

TUGAS AKHIR - KS 141501

**KLASIFIKASI PENYAKIT PARU DENGAN METODE
ARTIFICIAL NEURAL NETWORK (ANN) (STUDI KASUS:
RSUD KERTOSONO)**

***CLASSIFICATION OF LUNG DISEASE WITH ARTIFICIAL
NEURAL NETWORK (ANN) METHOD (CASE STUDY:
RSUD KERTOSONO)***

**NURITA DAMAYANTI
NRP 5213 100 083**

**Dosen Pembimbing:
Edwin Riksakomara, S.Kom, MT**

**DEPARTEMEN SISTEM INFORMASI
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember
Surabaya 2017**



ITS
Institut
Teknologi
Sepuluh Nopember

TUGAS AKHIR - KS 141501

KLASIFIKASI PENYAKIT PARU DENGAN METODE ARTIFICIAL NEURAL NETWORK (ANN) (STUDI KASUS: RSUD KERTOSONO)

NURITA DAMAYANTI
NRP 5213 100 083

Dosen Pembimbing
Edwin Riksakomara, S.Kom, MT



ITS
Institut
Teknologi
Sepuluh Nopember

FINAL PROJECT - KS 141501

***CLASSIFICATION OF LUNG DISEASE WITH
ARTIFICIAL NEURAL NETWORK (ANN)
METHOD (CASE STUDY: RSUD KERTOSOONO)***

NURITA DAMAYANTI
NRP 5213 100 083

Supervisor
Edwin Riksakomara, S.Kom, MT

LEMBAR PENGESAHAN

KLASIFIKASI PENYAKIT PARU DENGAN METODE ARTIFICIAL NEURAL NETWORK (ANN) (STUDI KASUS: RSUD KERTOSONO)

TUGAS AKHIR

Disusun untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada

Departemen Sistem Informasi
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember

Oleh:

NURITA DAMAYANTI
5213 100 083

Surabaya, 17 Juli 2017

**KEPALA
DEPARTEMEN SISTEM INFORMASI**


Dr. Ir. Aris Tjahyanto, M.Kom.
NIP 19650310 199102 1 001

LEMBAR PERSETUJUAN

KLASIFIKASI PENYAKIT PARU DENGAN METODE ARTIFICIAL NEURAL NETWORK (ANN) (STUDI KASUS: RSUD KERTOSONO)

TUGAS AKHIR

Disusun untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

pada

Departemen Sistem Informasi
Fakultas Teknologi Informasi

Institut Teknologi Sepuluh Nopember

Oleh :

NURITA DAMAYANTI

5213 100 083

Disetujui Tim Penguji : Tanggal Ujian : 5 Juli 2017

Periode Wisuda: September 2017

Edwin Riksakomara, S.Kom, M.T


(Pembimbing 1)

Wiwik Anggraeni, S.Si., M.Kom


(Penguji 1)

Ahmad Mukhlason, S.Kom, M.Sc, Ph.D


(Penguji 2)

KLASIFIKASI PENYAKIT PARU DENGAN METODE ARTIFICIAL NEURAL NETWORK (ANN) (STUDI KASUS: RSUD KERTOSONO)

Nama Mahasiswa : Nurita Damayanti
NRP : 5213 100 083
Departemen : SISTEM INFORMASI FTIF-ITS
Dosen Pembimbing : Edwin Riksakomara, S.Kom, M.T

ABSTRAK

Paru-paru merupakan salah satu organ pernapasan yang paling penting dalam tubuh manusia. Apabila paru-paru mengalami gangguan atau tidak berfungsi dengan baik, maka kebutuhan oksigen dalam tubuh juga tidak dapat terpenuhi dengan baik. Akibatnya, tubuh akan dipenuhi dengan gas karbondioksida (CO_2) yang bersifat racun bagi tubuh dan mengganggu setiap fungsi organ tubuh lainnya. Beberapa penyakit paru, seperti tuberculosi, PPOK, pneumonia, kanker paru dan bronkitis, telah menjadi penyebab kematian tertinggi di dunia, khususnya di Indonesia. Adapun permasalahan yang terdapat pada studi kasus RSUD Kertosono adalah belum adanya sistem diagnosa penyakit paru-paru untuk tenaga medis atau pasien secara real time di ruang poli paru sehingga pasien atau tenaga medis harus menunggu hasil uji laboratorium. Hasil uji laboratorium di RSUD Kertosono dapat diketahui sehari setelah uji laboratorium dilakukan karena banyaknya variabel yang dibutuhkan (rontgen organ paru-paru, darah, dahak). Selain uji laboratorium, tenaga medis juga dapat mendiagnosa jenis penyakit paru pasien dari beberapa variabel keluhan yang dirasakan secara langsung. Metode ANN dapat dengan cepat mengklasifikasikan pola dengan lebih dari 2 (dua) kelas target (multikelas). Oleh karena itu, penelitian ini bertujuan untuk membantu menyelesaikan permasalahan pada poli paru RSUD Kertosono dalam mendiagnosa penyakit paru-paru yang diderita pasien dengan menerapkan metode ANN. Hasil penelitian ini diharapkan dapat membantu tenaga medis dan pasien untuk lebih akurat

dan lebih cepat dalam mendiagnosa karena variabel yang digunakan tidak sebanyak yang dibutuhkan dalam uji laboratorium sehingga pasien dapat lebih cepat diobati dan ditangani.

Model ANN terbaik untuk mendiagnosa penyakit paru pada studi kasus penelitian ini adalah model feed-forward backpropagation neural network dengan 6 node pada input layer (jenis kelamin, umur, berat badan, diastole, batuk, sesak), 10 node pada 1 hidden layer, dan 4 node pada output layer (ppok, tuberculosis, bronkitis, pneumonia). Model ANN terbaik pada penelitian ini menghasilkan nilai akurasi sebesar 86.78%.

Kata Kunci: Klasifikasi, Diagnosa, Penyakit, Paru, Artificial Neural Network, Jaringan Syaraf Tiruan

CLASSIFICATION OF LUNG DISEASE WITH ARTIFICIAL NEURAL NETWORK (ANN) METHOD (CASE STUDY: RSUD KERTOSONO)

Name : Nurita Damayanti
NRP : 5213 100 083
Department : INFORMATION SYSTEMS FTIF-ITS
Supervisor : Edwin Riksakomara, S.Kom, M.T

ABSTRACT

The lungs are one of the most important respiratory organs in the human body. If the lungs are impaired or not functioning properly, then the need for oxygen in the body also can not be met properly. As a result, the body will be filled with carbon dioxide gas (CO₂) which is toxic to the body and interfere with every other organ function. Some lung diseases, such as tuberculosis, COPD, pneumonia, lung cancer and bronchitis, have been the leading cause of death in the world, especially in Indonesia. The problems that exist in Kertosono RSUD case study is the absence of a lung disease diagnosis system for medical personnel or patients in real time in the pulmonary poly room so that patients or medical personnel have to wait for laboratory test results. The result of laboratory test in RSUD Kertosono can be known one day after laboratory test is done because of the many variables needed (rontgen organ lungs, blood, sputum). In addition to laboratory tests, medical personnel can also diagnose the type of lung disease of patients from several directly perceived complaint variables.

The ANN method can quickly classify patterns with more than 2 (two) target classes (multikelas). Therefore, this study aims to help solve problems in the pulmonary tuberculosis RSUD Kertosono in diagnosing lung disease suffered by patients by applying the ANN method. The results of this study are expected to help medical personnel and patients to be more accurate and faster in diagnosing because the variables used are not as much

as needed in laboratory tests so that patients can be more quickly treated and handled.

The best ANN model for diagnosing lung disease in this case study study was a feed-forward backpropagation neural network model with 6 nodes on the input layer (gender, age, weight, diastole, cough, spasms), 10 nodes in 1 hidden layer, and 4 nodes on the output layer (Ppok, tuberculosis, bronchitis, pneumonia). The best ANN model in this study resulted in an accuracy of 86.78%.

Keywords: Classification, Diagnosis, Disease, Lung, Artificial Neural Network, Artificial Neural Network

KATA PENGANTAR

Puji syukur penulis panjatkan atas kehadiran Tuhan Yang Maha Esa atas segala berkat dan rahmat-Nya lah penulis dapat menyelesaikan buku tugas akhir dengan judul “**KLASIFIKASI PENYAKIT PARU DENGAN METODE ARTIFICIAL NEURAL NETWORK (ANN) (STUDI KASUS: RSUD KERTOSONO)**” yang merupakan salah satu syarat kelulusan pada Departemen Sistem Informasi, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember Surabaya.

Secara khusus penulis akan menyampaikan ucapan terima kasih yang sedalam-dalamnya kepada:

1. Ibu Sunarti dan Bapak Jasmin selaku kedua orang tua, serta segenap keluarga penulis yang selalu memberikan dukungan dan motivasi. Terima kasih atas doa dan dukungannya yang terus mengalir tiada henti.
2. Bapak Edwin Riksakomara, S.Kom, MT selaku dosen pembimbing dengan penuh keikhlasan dan dedikasi tinggi telah membimbing penulis dalam mengerjakan tugas akhir ini hingga selesai. Terima kasih atas kesediaan, waktu, dan ilmu yang telah diberikan.
3. Ibu Wiwik Anggraeni S.Si., M.Kom dan Bapak Faisal Mahananto, S.Kom, M.Eng, Ph.D selaku dosen penguji yang selalu memberikan saran dan masukan guna kebaikan Tugas Akhir ini.
4. Untuk seluruh teman-teman SI 2013 (Beltranis) dan Jama'ah Masjid Manarul Ilmi (JMMI) yang kerap menjadi tempat berbagi keluh kesah dan dukungan selama kuliah.
5. Seluruh dosen pengajar, staff, dan karyawan di Departemen Sistem Informasi FTIF ITS Surabaya yang telah memberikan ilmu dan bantuan kepada penulis selama ini.

6. Serta semua pihak yang telah membantu dalam pengerjaan Tugas Akhir ini yang belum mampu penulis sebutkan diatas.

Terima kasih atas segala bantuan, dukungan, serta doanya. Semoga Allah SWT senantiasa memberikan kesehatan, keselamatan, karunia dan nikmat-Nya.

Penulis pun ingin memohon maaf karena Penulis menyadari bahwa Tugas Akhir ini masih belum sempurna dengan segala kekurangan di dalamnya. Selain itu penulis bersedia menerima kritik dan saran terkait dengan Tugas Akhir ini. Semoga Tugas Akhir ini dapat bermanfaat bagi seluruh pembaca.

Surabaya, Juni 2017

DAFTAR ISI

ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan	4
1.5 Manfaat	4
1.6 Relevansi	5
BAB II TINJAUAN PUSTAKA	7
2.1 Penelitian Sebelumnya	7
2.2 Dasar Teori	13
2.2.1 Knowledge Discovery in Databases	13
2.2.2 Interpolasi	14
2.2.3 Uji Normalitas	16
2.2.4 Uji Homogenitas	17
2.2.5 Uji Analysis of Variances	18
2.2.6 Uji Korelasi	20
2.2.7 Penyakit Paru-paru	21
2.2.8 Klasifikasi	23
2.2.9 Artificial Neural Network	24
2.2.10 Recall, Precision dan Accuracy	26
BAB III METODOLOGI PENELITIAN	28
3.1 Diagram Metodologi	28
3.2 Uraian Metodologi	29
3.2.1 Analisa Permasalahan	29
3.2.2 Studi Pustaka	29
3.2.3 Pengumpulan Data dan Informasi	29
3.2.4 Pembuatan Model	30
3.2.5 Training dan Uji Coba Model	30
3.2.6 Validasi Model dan Analisa	30

BAB IV PERANCANGAN	32
4.1 Persiapan Data	32
4.1.1 Uji Normalitas	34
4.1.2 Uji Homogenitas.....	48
4.1.3 Analysis of Variances (Uji F).....	52
4.1.4 Uji Korelasi	59
4.1.5 Karakteristik Data.....	60
4.2 Penentuan Model Neural Nework.....	64
4.3 Penentuan Struktur Model ANN.....	65
4.3.1 Penentuan Input Layer.....	66
4.3.2 Penentuan Output Layer	66
4.3.3 Penentuan Hidden Layer	68
4.4 Penentuan Parameter	70
4.4.1 Fungsi Aktivasi	70
4.4.2 Fungsi Pelatihan	72
4.4.3 Jumlah Epoch	74
4.4.4 Fungsi Performa	75
4.4.5 Fungsi Pembelajaran	75
BAB V IMPLEMENTASI.....	77
5.1 Pemodelan ANN dengan MATLAB.....	77
5.1.1 Memasukkan Data.....	77
5.1.2 Membuat Model ANN	77
5.1.3 Melakukan Training	78
5.1.4 Melakukan Testing	78
5.2 Validasi Model.....	79
BAB VI HASIL DAN PEMBAHASAN	81
6.1 Lingkungan Uji Coba	81
6.2 Implementasi Model 1 Node Output	81
6.2.1 Implementasi 5 Variabel Input	82
6.2.2 Implementasi 6 Variabel Input	83
6.2.3 Implementasi 7 Variabel Input	85
6.3 Implementasi Model 4 Node Output.....	86
6.3.1 Implementasi 5 Variabel Input.....	87
6.3.2 Implementasi 6 Variabel Input.....	88
6.3.3 Implementasi 7 Variabel Input.....	89

6.4	Perbandingan Recall, Precision, dan Accuracy ...	91
6.5	Kesimpulan Hasil Percobaan.....	95
BAB VII KESIMPULAN DAN SARAN		98
7.1	Kesimpulan	98
7.2	Saran	98
DAFTAR PUSTAKA		100
BIODATA PENULIS		102
A.	LAMPIRAN A DATA MENTAH (7 VARIABEL AWAL)	A-1
B.	LAMPIRAN B TABEL KONTINGENSI.....	B-1
C.	LAMPIRAN C DOKUMENTASI APLIKASI....	3

DAFTAR GAMBAR

Gambar 1 Struktur ANN [10].....	25
Gambar 2 Single layer network [10]	26
Gambar 3 Three-layer network [10].....	26
Gambar 4 Histogram persebaran data variabel jenis kelamin	36
Gambar 5 Histogram persebaran data variabel umur	38
Gambar 6 Histogram persebaran data variabel berat badan...	40
Gambar 7 Histogram persebaran data variabel sistole	42
Gambar 8 Histogram persebaran data variabel diastole	44
Gambar 9 Histogram persebaran data variabel batuk	46
Gambar 10 Histogram persebaran data variabel sesak.....	48
Gambar 11 Struktur ANN dengan 1 variabel input.....	67
Gambar 12 Struktur ANN dengan 4 variabel input.....	68
Gambar 13 Struktur Model 5 Node Input dan 1 Node Output	82
Gambar 14 Grafik Performa Model 5 Node Input dan 1 Node Output.....	83
Gambar 15 Struktur Model 6 Node Input dan 1 Node Output	84
Gambar 16 Grafik Performa Model 6 Node Input dan 1 Node Output.....	85
Gambar 17 Struktur Model 7 Node Input dan 1 Node Output	86
Gambar 18 Grafik Performa Model 7 Node Input dan 1 Node Output.....	86
Gambar 19 Struktur Model 7 Node Input dan 4 Node Output	87
Gambar 20 Grafik Performa Model 7 Node Input dan 4 Node Output.....	88
Gambar 21 Struktur Model 6 Node Input dan 4 Node Output	88
Gambar 22 Grafik Performa Model 6 Node Input dan 4 Node Output.....	89

Gambar 23 Struktur Model 7 Node Input dan 4 Node Output	90
Gambar 24 Grafik Performa Model 7 Node Input dan 4 Node Output.....	91
Gambar 25 Grafik nilai MSE	93
Gambar 26 Grafik nilai akurasi	94
Gambar 27 Grafik nilai precision.....	94
Gambar 28 Grafik nilai recall.....	94
Gambar 29 Tampilan Aplikasi di Matlab.....	3

DAFTAR TABEL

Tabel 1 Penelitian Terkait	7
Tabel 2 Kriteria data berdistribusi normal	17
Tabel 3 Rumus uji anova.....	19
Tabel 4 Besaran nilai korelasi	21
Tabel 5 Metodologi Penelitian	28
Tabel 6 Pembagian Data Training dan Testing	32
Tabel 7 Nilai korelasi antar variabel diagnosa	33
Tabel 8 Hasil uji normalitas variabel jenis kelamin	35
Tabel 9 Hasil uji normalitas variabel umur	37
Tabel 10 Hasil uji normalitas variabel berat badan	39
Tabel 11 Hasil uji normalitas variabel sistole	41
Tabel 12 Hasil uji normalitas variabel diastole	42
Tabel 13 Hasil uji normalitas variabel batuk.....	44
Tabel 14 Hasil uji normalitas variabel sesak.....	46
Tabel 15 Kesimpulan hasil uji normalitas.....	48
Tabel 16 Deskripsi statistik data variabel umur	49
Tabel 17 Hasil uji homogenitas variabel umur	49
Tabel 18 Deskripsi statistik data variabel berat badan	50
Tabel 19 Hasil uji homogenitas variabel berat badan	50
Tabel 20 Deskripsi statistik data variabel sistole	51
Tabel 21 Hasil uji homogenitas variabel sistole.....	51
Tabel 22 Deskripsi statistik data variabel diastole	51
Tabel 23 Hasil uji homogenitas variabel diastole	52
Tabel 24 kesimpulan hasil uji homogenitas	52
Tabel 25 Hasil uji Kruskal Wallis variabel jenis kelamin.....	53
Tabel 26 Hasil uji anova variabel umur	54
Tabel 27 Perbandingan deskripsi umur pada setiap jenis penyakit.....	54
Tabel 28 Hasil uji anova variabel berat badan	54
Tabel 29 Perbandingan deskripsi berat badan pada setiap jenis penyakit.....	55
Tabel 30 Hasil uji anova variabel sistole	55

Tabel 31 Perbandingan deskripsi sistole pada setiap jenis penyakit.....	56
Tabel 32 Hasil uji anova variabel diastole	56
Tabel 33 Perbandingan deskripsi diastole pada setiap jenis penyakit.....	57
Tabel 34 Hasil uji Kruskal Wallis variabel batuk	57
Tabel 35 Hasil uji Kruskal Wallis variabel sesak.....	58
Tabel 36 Kesimpulan hasil uji anova dan kruskal wallis	58
Tabel 37 Hasil Uji Korelasi.....	59
Tabel 38 Statistika deskriptif variabel titik embun	60
Tabel 39 Rancangan Struktur Model ANN.....	65
Tabel 40 Struktur Model ANN dengan Fungsi Pembelajaran Trainlm.....	68
Tabel 41 Struktur Model ANN dengan Fungsi Pembelajaran Trainbr.....	69
Tabel 42 Variasi Fungsi Aktivasi pada 1 Node Output.....	71
Tabel 43 Variasi Fungsi Aktivasi pada 1 Node Output.....	71
Tabel 44 Fungsi Pembelajaran Neural Network	72
Tabel 45 Penentuan Fungsi Pembelajaran	74
Tabel 46 Klasifikasi nilai keluaran dengan 1 node pada output layer.....	79
Tabel 47 Contoh penentuan diagnosa dengan 4 node pada output layer.....	79
Tabel 48 Lingkungan perangkat keras yang digunakan	81
Tabel 49 Lingkungan perangkat lunak yang digunakan	81
Tabel 50 Struktur model ANN hasil percobaan	91
Tabel 51 Perbandingan Recall, Precision, dan Accuracy 1 Node Output.....	92
Tabel 52 Perbandingan Recall, Precision, dan Accuracy 4 Node Output.....	92
Tabel 53 Data mentah penyakit ppok (7 variabel)	A-1
Tabel 54 Data mentah penyakit tuberculosis (7 variabel)....	A-3
Tabel 55 Data mentah penyakit bronkitis (7 variabel)	A-15
Tabel 56 Data mentah penyakit pneumonia (7 variabel) ...	A-18
Tabel 57 Tabel Kontingensi 5 Node Input 1 Node Output ..	B-1

Tabel 58 Tabel Kontingensi 6 Node Input 1 Node Output ..B-1
Tabel 59 Tabel Kontingensi 7 Node Input 1 Node Output ..B-1
Tabel 60 Tabel Kontingensi 5 Node Input 4 Node Output ..B-2
Tabel 61 Tabel Kontingensi 6 Node Input 4 Node Output ..B-2
Tabel 62 Tabel Kontingensi 7 Node Input 4 Node Output ..B-2

BAB I

PENDAHULUAN

Pada bab ini, akan menjelaskan mengenai latar belakang masalah, perumusan masalah, batasan masalah, tujuan tugas akhir, manfaat tugas akhir, serta relevansi penelitian tugas akhir dengan bidang keilmuan sistem informasi.

1.1 Latar Belakang

Selama hidupnya, manusia akan selalu melakukan proses pernapasan. Pernapasan merupakan proses yang sangat vital bagi keberlangsungan hidup manusia. Proses pernapasan mengambil oksigen yang dibutuhkan oleh setiap sel pada tubuh, serta melepaskan karbon dioksida yang bersifat racun bagi tubuh. Untuk melakukan proses pernapasan, diperlukan organ-organ khusus. Organ-organ tersebut memiliki struktur dan fungsi yang berbeda. Apabila ada salah satu dari organ tersebut tidak berfungsi atau sakit, maka akan mengganggu kinerja organ pernapasan lainnya dan menghambat proses pernapasan, bahkan akibat paling buruk bisa mengakibatkan kematian.

Salah satu organ pernapasan yang memegang peranan paling penting adalah paru-paru. Penyakit yang dapat menyerang paru-paru, diantaranya tuberculosis (TB), radang paru-paru (pneumonia), Paru Paru Obstruktif Kronik (PPOK), bronkitis, asma, dan kanker paru-paru.

Penyakit tuberculosis (TB) merupakan penyakit menular yang masih menjadi masalah kesehatan masyarakat dan salah satu penyebab kematian. Tuberculosis menyebar melalui udara ketika seseorang dengan infeksi TB aktif batuk, bersin, atau menyebarkan butiran ludah mereka melalui udara. Orang dengan sistem kekebalan tubuh lemah memiliki risiko yang jauh lebih besar untuk terjangkit TB. TB merupakan salah satu dari 10 kasus utama kematian di seluruh dunia. Peringkat tersebut di atas kasus kematian akibat HIV dan malaria. Diperkirakan sekitar sepertiga penduduk dunia terinfeksi bakteri tuberculosis (TB). Pada tahun 2014, World Health

Organization (WHO) mencanangkan “End TB Strategy” sebagai bentuk usaha untuk mencegah, mendeteksi, dan mengobati TB hingga 90% pada tahun 2030. Pada tahun 2015, terdapat sekitar 10.4 juta kasus TB di seluruh dunia. 60% dari total tersebut sebagian besar berasal dari negara India, Indonesia, China, Nigeria, Pakistan Afrika Selatan. [1]

Pneumonia terhitung 16% penyebab kematian anak-anak di bawah 5 tahun pada tahun 2015. Pneumonia disebabkan oleh virus, bakteri, dan jamur. Pneumonia dapat dicegah dengan imunisasi, nutrisi, dan faktor lingkungan yang sehat. Diperkirakan hanya sepertiga anak-anak yang menderita pneumonia yang mendapatkan antibiotik yang mereka butuhkan. [1]

Penyakit Paru Obstruktif Kronis (PPOK) merupakan istilah yang digunakan untuk menggambarkan penyakit paru yang sangat buruk, seperti asma jangka panjang (tidak ada perubahan), bronkitis menahun atau kronis, emfisema (kondisi kantung udara di paru-paru mengalami kerusakan yang sangat buruk) [6]. Sebagian besar penderita PPOK dikarenakan tembakau dari rokok (baik perokok aktif maupun pasif). PPOK ini juga merupakan salah satu penyakit dari 10 kasus penyakit dengan jumlah kasus kematian terbesar di dunia sejak tahun 2000. 3,2 juta penduduk dunia meninggal dunia akibat PPOK, ketika kanker paru (termasuk kanker trachea dan bronkus) menyebabkan 1,7 juta penduduk meninggal dunia pada tahun 2015. PPOK tidak dapat disembuhkan, tetapi pengobatan pada penderita dapat meringankan gejala, meningkatkan kualitas hidup dan risiko kematian. [1]

Bronkitis merupakan infeksi akibat bakteri atau virus pada saluran pernapasan utama (bronkus) yang menyebabkan terjadinya peradangan pada saluran tersebut. Bronkitis yang berulang (kronis) dapat menjadi salah satu gejala dari PPOK [5]. Pengenalan dan pengobatan lebih dini penyakit bronkitis masih harus terus dilakukan sebab penyakit tersebut masih menjadi salah satu penyakit dari 10 penyakit dengan jumlah

kematian terbesar di dunia sejak tahun 2000, bahkan mengalami peningkatan di tahun 2015. [1]

Berdasarkan uraian singkat mengenai fakta dan dampak penyakit yang menyerang organ paru-paru tersebut, dapat disimpulkan bahwa penyakit paru-paru merupakan penyakit yang sangat vital untuk dikenali dan dicegah lebih dini.

Dewasa ini, seiring berkembangnya ilmu pengetahuan di bidang teknologi (komputasional), hampir setiap pekerjaan manusia dapat dilakukan oleh sistem dengan kecerdasan buatan. Salah satu sistem dengan kecerdasan buatan yang berhasil dikembangkan saat ini yaitu proses pembelajaran mesin (komputer). Suatu sistem dengan kecerdasan buatan dapat dikatakan mengalami pembelajaran mesin apabila performa dari program mesin tersebut mengalami peningkatan seiring dengan semakin banyaknya pengalaman dalam pengerjaan tugasnya. Salah satu model pembelajaran mesin yang meniru cara kerja sistem saraf biologis manusia adalah sistem jaringan syaraf tiruan (Artificial Neural Network).

Proses diagnosa penyakit bagi tenaga medis mempunyai peranan yang sangat penting, tidak hanya tingkat akurasi yang tinggi, tetapi juga tingkat kecepatan dalam mendiagnosis penyakit yang diderita pasien. Apabila tenaga medis mengalami kesalahan atau keterlambatan dalam mendiagnosis penyakit pasien, maka bentuk pengobatan dan penanganan terhadap pasien juga akan salah. Oleh sebab itu, dalam proses diagnosis penyakit terutama penyakit yang berhubungan dengan organ pernapasan, diperlukan diagnosa dengan tingkat akurasi yang tinggi dan proses diagnosis yang cepat [13].

Artificial Neural network (ANN) merupakan salah satu sistem kecerdasan buatan yang mampu menganalisis data yang kompleks dan memiliki hubungan antar variabel yang tidak terdefinisi atau non-linier [9]. Pada penelitian sebelumnya, metode ANN ini telah banyak diterapkan dalam bidang kesehatan untuk mendiagnosa penyakit, seperti tuberculosi dengan tingkat akurasi 95.08%, pneumonia dengan tingkat

akurasi 82.8%, PPOK dengan tingkat akurasi 96.08%, asma dengan tingkat akurasi 90% [2].

Oleh sebab itu, pada penelitian tugas akhir ini, akan menggunakan metode Artificial Neural network (ANN) untuk mendiagnosa jenis penyakit paru-paru khususnya tuberculosis, pneumonia, PPOK dan bronkitis. Sehingga dapat membantu tenaga medis untuk mendiagnosa penyakit paru dengan tingkat akurasi yang tinggi dan dalam waktu yang lebih cepat.

1.2 Rumusan Masalah

1. Bagaimana mendiagnosa jenis penyakit paru-paru dengan menggunakan metode Artificial Neural Network (ANN)?
2. Bagaimana tingkat keakuratan metode Artificial Neural Network (ANN) dalam mendiagnosa jenis penyakit paru-paru?

1.3 Batasan Masalah

Batasan permasalahan dalam tugas akhir ini adalah

1. Data yang digunakan dalam penelitian ini diperoleh dari RSUD Kertosono, Nganjuk.
2. Data diambil dari pasien rawat jalan.
3. Data diambil pada periode Januari 2014-Desember 2016.
4. Jenis penyakit paru yang digunakan hanya 4 jenis (tuberculosis, pneumonia, PPOK, bronkitis).

1.4 Tujuan

1. Untuk mendiagnosa jenis penyakit paru-paru dengan menggunakan metode Artificial Neural Network (ANN).
2. Untuk mendeskripsikan tingkat keakuratan metode Artificial Neural Network (ANN) dalam mendiagnosa jenis penyakit paru-paru.

1.5 Manfaat

1. Manfaat bagi peneliti:

Menambah pengetahuan dan pengalaman dalam menggunakan metode Artificial Neural Network (ANN) untuk mendiagnosis jenis penyakit paru.

2. Manfaat untuk instansi kesehatan/ahli medis:
Membantu dokter untuk mendiagnosis jenis penyakit paru.
3. Manfaat untuk masyarakat:
Memudahkan masyarakat untuk melakukan diagnosis penyakit paru secara mandiri, sehingga dapat segera diobati dan ditangani.
4. Manfaat untuk universitas:
Menambah referensi tentang penggunaan metode Artificial Neural Network (ANN) dalam mendiagnosis jenis penyakit paru.

1.6 Relevansi

Hasil dari penelitian tugas akhir ini difokuskan pada keakuratan hasil diagnosis yang didapat. Penelitian tugas akhir ini termasuk dalam mata kuliah sistem Sistem Cerdas, Statistik, Sistem Pendukung Keputusan, Teknik Peramalan, Penggalian Data dan Analitika Bisnis, serta termasuk dalam topik dalam laboratorium Rekayasa Data dan Intelegensia Bisnis di Jurusan Sistem Informasi.

Halaman ini sengaja dikosongkan

BAB II

TINJAUAN PUSTAKA

Pada bab ini, akan menjelaskan mengenai penelitian sebelumnya dan dasar teori yang digunakan sebagai bahan penunjang dalam penyusunan tugas akhir ini..

2.1 Penelitian Sebelumnya

Tabel 1 memuat daftar penelitian yang telah dilakukan sebelumnya yang mendasari penelitian tugas akhir ini.

Tabel 1 Penelitian Terkait

No.	Penelitian Sebelumnya	
1	Penulis; Tahun	Joao B. de O. Souza Filho, Jose manoel de Seixas, Rafael Galliez, Basilio de Braganca Pereira, Fernanda C. de Q Mello, Alcione Miranda dos Santos, Afranio Lineu Kritski (Federal University of Rio de Janeiro, Brazil); 2016.
	Judul Penelitian	<i>A screening system for smear-negative pulmonary tuberculosis using artificial neural networks.</i>
	Tujuan Penelitian	Membuat sebuah sistem penyeleksi dan penilaian risiko untuk penyakit tuberculosis paru-paru smear-negative dengan menggunakan ANN berdasarkan gejala yang dialami pasien.
	Keterangan	<ul style="list-style-type: none"> • Metode: CART, Multivariate logistic regression, SVM polynomial, SVM RBF, SVM linier, <i>Multilayer perceptron network</i>. • Tingkat sensitivitas, akurasi, dan AUC paling tinggi didapat dengan metode MLP (100%, 88%, 0.918), sedangkan tingkat spesifisitas paling tinggi didapat dengan metode <i>Multivariate logistics regression</i> (85%). • Terdiri dari 12 input nodes (umur, batuk, demam, haemoptysis, anorexia,

No.	Penelitian Sebelumnya	
		<p>penurunan berat badan, HIV, night sweats, dyspnoea, merokok, ekstra paru TB, riwayat pengobatan); 15 hidden neurons; 1 output neuron (kasus negatif dan positif).</p> <ul style="list-style-type: none"> • Sample: 136 data (75% data training dan 25% data testing).
2	Penulis; Tahun	Yongjun Wu, Yiming Wu, Jing Wang, Zhen Yan, Lingbo Qu, Bingren Xiang, Yiguo Zhang (Zhengzhou University, China); 2011.
	Judul Penelitian	<i>An optimal tumor marker group-coupled artificial neural network for diagnosis of lung cancer.</i>
	Tujuan Penelitian	Mengombinasikan tumor-marker dengan ANN untuk meningkatkan efisiensi (sensitivitas dan spesifisitas) dalam mendiagnosa penyakit kanker paru-paru.
	Keterangan	<ul style="list-style-type: none"> • Metode: ANN. • Tingkat spesifisitas meningkat dari 72% menjadi 100%, dan tingkat akurasi meningkat dari 71.4% menjadi 92.8%. • Sample: 140 data (100 data training dan 40 data testing). • 50 data positif kanker paru, 40 data positif menderita penyakit paru jenis lainnya, 50 data kondisi paru normal.
3	Penulis; Tahun	Orhan Er, Nejat Yumusak, Feyzullah Temurtas (Bozok University, Turkey); 2010.
	Judul Penelitian	<i>Chest disease diagnosis using artificial neural networks.</i>
	Tujuan Penelitian	Membandingkan tingkat akurasi hasil diagnosa 5 jenis penyakit paru-paru dari beberapa struktur metode ANN yang berbeda.

No.	Penelitian Sebelumnya	
	Keterangan	<ul style="list-style-type: none"> • Metode: <i>Multilayer neural network, probabilistic neural network, learning vector quantization neural network, generalized regression neural network, radial basis function neural network.</i> • Tingkat akurasi tertinggi didapat dengan menggunakan metode probabilistic neural network (92.16%). • Terdiri dari 38 input layer; 1 hingga 2 hidden layer; 6 output layer (tuberculosis, PPOK, pneumonia, asma, kanker paru, dan normal). • Sample: 357 data (50 data tuberculosis, 71 data PPOK, 60 data pneumonia, 44 data asma, 32 data kanker paru, 100 data normal).
4	Penulis; Tahun	Ratnaningtyas Widya Purnamasari, Dwijanto, Endang Sugiharti (Jurusan Matematika, FMIPA, Universitas Negeri Semarang); 2013.
	Judul Penelitian	Implementasi Jaringan Syaraf Tiruan Backpropagation sebagai Sistem Deteksi Penyakit Tuberculosis (TBC).
	Tujuan Penelitian	Mengetahui arsitektur Jaringan Syaraf Tiruan (JST) yang maksimal dan mengetahui besarnya kemampuan JST dalam mengklasifikasikan data.
	Keterangan	<ul style="list-style-type: none"> • Metode: Artificial Neural Network (ANN) backpropagation. • Sistem dibuat dengan menggunakan perangkat lunak Matlab R2009a. • Sistem mendiagnosa 1 (satu) jenis penyakit paru, yaitu TBC. • Terdiri dari 6 input layer, yaitu batuk, sesak napas, nyeri dada, demam, penurunan nafsu makan, dan badan

No.	Penelitian Sebelumnya	
		<p>lemah; 1 hidden layer; 2 output layer, yaitu positif TBC dan negatif TBC.</p> <ul style="list-style-type: none"> • Tingkat akurasi: 100%, learning rate: 0.5, node: 100, nilai toleransi error: 0.001, jumlah epoch: 1000. • Sample: 80 data training, dan 20 data testing.
5	Penulis; Tahun	Dahriani Hakim Tanjung (STMIK Potensi Utama Medan); 2015
	Judul Penelitian	Jaringan Saraf Tiruan dengan Backpropagation untuk Memprediksi Penyakit Asma.
	Tujuan Penelitian	Memprediksi penyakit asma menggunakan teknik pengenalan pola yaitu jaringan saraf tiruan dengan metode backpropagation.
	Keterangan	<ul style="list-style-type: none"> • Metode: Artificial Neural Network (ANN) backpropagation. • Sistem dibuat dengan menggunakan perangkat lunak Matlab. • Sistem mendiagnosa 1 (satu) jenis penyakit paru, yaitu asma. • Terdiri dari 18 input layer; 8 hidden layer; 4 output layer. • Tingkat akurasi: 100%, learning rate: 0.5, nilai toleransi error: 0.001, jumlah epoch: 4707, MSE: 0.001. • Sample: 20 data, yang terdiri dari 6 data asma akut persisten ringan, 4 data asma akut persisten sedang, 6 data asma akut persisten berat, 4 data asma akut mengancam jiwa.
6	Penulis; Tahun	Ali A. El-Solh, MD; Chiu-Bin Hsiao, MD; Susan Goodnough, RN; Joseph Serghani, MD; and Brydon J. B. Grant, MD, FCCP; 1997
	Judul Penelitian	<i>Predicting Active Pulmonary Tuberculosis Using an Artificial Neural Network</i>

No.	Penelitian Sebelumnya	
	Tujuan Penelitian	Untuk memprediksi penyakit TB paru secara <i>real time</i> dengan mengembangkan metode ANN.
	Keterangan	<ul style="list-style-type: none"> • Metode: General Regression Neural Network • Sample: 682 (563 data training, 119 data validation) • Variabel input terdiri dari demografis (umur, Cluster of Diffrentiation 4, diabetes mellitus, HIV, Purified protein derivative), gejala konstitusional (sakit dada, penurunan berat badan, batuk, keringat malam, demam, sesak nafas), radiografis (infiltrate lobus atas, infiltrate lobus bawah, cavity lobus atas, cavity lobus bawah, adenopathy, unilateral pleural effusion, bilateral pleural effusion, pleural thickening, military pattern, normal). Dengan output yaitu negatif dan positif TB paru. Dan jumlah hidden layer yang lebih dari jumlah variabel input. • Sensitivitas dengan metode ANN mengalami kenaikan dari 64% menjadi 100% (95% CI), dan spesifisitas dari 79% menjadi 69% (95% CI).
7	Penulis; Tahun	Paul S. Heckerling, Ben S. Gerber, Thomas G. Tape, Robert S. Wigton; 2003
	Judul Penelitian	<i>Use of genetic algorithms for neural networks to predict community-acquired pneumonia</i>
	Tujuan Penelitian	Menggunakan algoritma genetika untuk menentukan jumlah hidden layer, konektivitas, dan training parameter yang optimal pada metode ANN yang digunakan untuk memprediksikan akuisisi data pneumonia.

No.	Penelitian Sebelumnya	
	Keterangan	<ul style="list-style-type: none"> • Metode: <i>Feed-forward backpropagation</i> ANN • Sample: 1044 data training, 116 data testing. • Model terdiri dari 35 variabel input yang terdiri dari data demografis (umur, jenis kelamin, ras), gejala (batuk, dahak, demam, dingin, nyeri dada karena pleuritis, nyeri dada tanpa pleuritis, sesak napas, mengi, ortopnea, sesak napas nocturnal paroxysmal), sign (suhu, tingkat pernapasan, denyut nadi, status gangguan mental, splinting, sianosis, perkusi pekak, rales, ronki, mengi, suara napas yang dikeluarkan, suara napas bronkial, egophony, pleural friction rub, penurunan toraks ekspansi), komorbiditas (asma, PPOK, penyakit paru lainnya, gagal jantung kongestif, penyakit immunocompromising, dementia, kondisi kormobid lainnya); 9 nodes pada hidden layer, serta output: aktif dan negatif pneumonia. • Akurasi 85% ROC.
8	Penulis; Tahun	Hanif, N. H. H. M., Lan, W. H., Daud, H. B., & Ahmad, J. (Malaysia); 2009.
	Judul Penelitian	<i>Classification of control measures for asthma using artificial neural networks</i>
	Tujuan Penelitian	Untuk mengklasifikasikan tindakan pengendalian asma menggunakan jaringan syaraf tiruan.
	Keterangan	<ul style="list-style-type: none"> • Metode: <i>backpropagation network</i>, <i>radial basis function network</i>. • Akurasi 90%.
9	Penulis; Tahun	Er O, Temurtas F (Sakarya University, Turkey); 2008

No.	Penelitian Sebelumnya	
	Judul Penelitian	<i>A study on chronic obstructive pulmonary disease diagnosis using multilayer neural networks.</i>
	Tujuan Penelitian	Untuk mendiagnosa penyakit PPOK dengan menggunakan multilayer neural networks
	Keterangan	<ul style="list-style-type: none"> • Metode: MNN Backpropagation • Terdiri dari 1-2 hidden layer. • Akurasi: 96.08%.

2.2 Dasar Teori

2.2.1 Knowledge Discovery in Databases

Knowledge Discovery in Databases (KDD) merupakan ilmu pengetahuan yang digunakan dalam pengolahan sumber data (database) dalam *data mining*. Terdapat 3 proses utama dalam KDD, yaitu:

2.2.1.1 Preprocessing

Data yang terdapat dalam database sebagian besar adalah data yang kompleks (jumlahnya sangat besar), disajikan berupa data mentah yang didapatkan langsung dari user dan belum dilakukan pengolahan lebih lanjut untuk menangani data yang *error*, tidak lengkap, *not available* (kosong). Proses ini meliputi *data cleaning*, *data integration*, *data transformation*, dan *data reduction*.

Proses *data cleaning* digunakan untuk menangani nilai yang kosong, nilai yang *error* atau *outlier* (*noisy data*), dan data yang tidak konsisten. Integrasi data yaitu mengintegrasikan berbagai sumber data untuk memperjelas dan mendukung data utama. Untuk mengatasi nilai data yang kosong, dapat ditangani dengan menghapus baris, mengisi nilai yang kosong dengan nilai konstanta global atau nilai rata-rata secara manual. Untuk menangani data yang *error* atau *outlier*, dapat dihaluskan

dengan metode *binning*, *clustering*, inspeksi komputer, dan regresi. Sedangkan untuk menangani data yang tidak konsisten dapat dicek melalui hubungan ketergantungan (*dependencies*) antar *attributes*.

Transformasi data yaitu mentransformasikan atau mengonsolidasikan data menjadi bentuk yang sesuai untuk *data mining*. Transformasi data dapat dilakukan dengan normalisasi data, *smoothing*, agregasi, atau generalisasi data.

Sedangkan reduksi data yaitu mengurangi jumlah yang besar dan kompleks dari sumber data untuk mengurangi waktu yang lama dalam proses *data mining*. Pengurangan data ini dilakukan bukan untuk mengurangi integritas dan kualitas data asli, melainkan mengurangi jumlah dimensi atau *attributes*.

2.2.1.2 Proses Data Mining

Proses data mining merupakan proses utama dalam KDD, yaitu menerapkan algoritma data mining untuk mendapatkan pengetahuan dari sumber data.

2.2.1.3 Post Processing

Hasil yang diperoleh dari proses *data mining* selanjutnya akan dievaluasi pada *post processing*. Selain itu, pengetahuan yang didapatkan +dari data mining, dapat divisualisasikan dan didokumentasikan untuk *end user*.

2.2.2 Interpolasi

Interpolasi merupakan salah satu teknik dalam analisis numerik yang digunakan untuk menghasilkan titik-titik (nilai-nilai) baru yang belum diketahui sebelumnya dari suatu set data diskrit yang sudah diketahui sebelumnya. Suatu set data (variabel) dianggap membentuk suatu grafik (fungsi) terhadap variabel lain yang memiliki korelasi terhadap variabel tersebut. Sehingga dari gambaran grafik (fungsi) tersebut dapat diperkirakan nilai-nilai kosong yang terdapat pada garis tersebut. [25]

2.2.2.1 Jenis-jenis Interpolasi

Jenis-jenis interpolasi dibedakan berdasarkan jumlah derajat (orde). Semakin tinggi jumlah orde yang digunakan, semakin tinggi tingkat keakuratan nilai (titik) yang dihasilkan. [25]

2.2.2.1.1 Interpolasi Linier

Interpolasi linier adalah interpolasi paling sederhana, yaitu dengan menggunakan hubungan berupa garis lurus antara dua titik. Interpolasi linier ini merupakan interpolasi dengan derajat (orde) 1. Hubungan antara 2 (dua) titik tersebut direpresentasikan ke dalam fungsi garis lurus (linier). Titik (nilai) baru pada garis tersebut dapat diketahui dengan memasukkan salah satu nilai koordinat (x,y) pada fungsi garis lurus tersebut. Fungsi garis lurus ditunjukkan pada persamaan (1). [25]

$$y = y_1 + \frac{(x-x_1)}{(x_2-x_1)}(y_2 - y_1) \dots (1)$$

2.2.2.1.2 Interpolasi Kuadrat

Interpolasi kuadrat adalah suatu teknik interpolasi yang digunakan untuk mendapatkan nilai (titik) yang belum diketahui dengan menggunakan 3 (tiga) titik terdekatnya. Interpolasi linier ini merupakan interpolasi dengan derajat (orde) 2. Misalkan 3 (tiga) titik tersebut adalah $P_0(x_0, y_0)$, $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, maka fungsi kuadrat dari ketiga titik tersebut dapat ditunjukkan pada persamaan (2). [25]

$$y = y_0 \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} + y_1 \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} + y_2 \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} \dots (2)$$

2.2.2.1.3 Interpolasi Polinom

Interpolasi polinom merupakan jenis interpolasi berderajat (orde) banyak atau lebih dari 2 orde. Interpolasi polinom memiliki 2 (dua) macam, yaitu interpolasi polynomial newton dan interpolasi polynomial lagrange. Interpolasi newton didasarkan pada formulasi beda hingga, dimana dibuat suatu polynomial dengan titik-titik data sebagai titik simpul.

Sedangkan interpolasi lagrange tidak menggunakan bentuk pembagian beda hingga. Bentuk interpolasi polynomial newton ditunjukkan pada persamaan (3). [25]

$$P_n(x) = C_0 + C_1(x - x_0) + C_2(x - x_0)(x - x_1) + \dots + C_n(x - x_0)(x - x_1) \dots (x - x_{n-1}) \dots (3)$$

dimana:

C_0, C_1, \dots, C_n merupakan suatu konstanta C_i dengan $i=0,1,\dots,n$.

Harga C_i dapat dirumuskan pada persamaan (4):

$$C_i = \frac{\Delta^i y}{(i!)h^i} \dots (4)$$

Sedangkan fungsi polinomial legrenge ditunjukkan pada persamaan (5).

$$P(x) = L_0(x)y_0 + L_1(x)y_1 + \dots + L_n(x)y_n \dots (5)$$

dimana:

$$L_n(x) = \frac{(x-x_{n+1})(x-x_{n+2})}{(x_n-x_{n+1})(x_n-x_{n+2})} \text{ dan } n=\text{jumlah orde.}$$

2.2.3 Uji Normalitas

Uji normalitas merupakan uji statistic yang digunakan untuk mengetahui persebaran satu set data pada masing-masing variabel, apakah berdistribusi normal atau tidak. Apabila data dinyatakan berdistribusi normal, data tersebut dapat dianggap mewakili populasi. Uji normalitas dilakukan dengan metode deskriptif dan analitis pada setiap variabel diagnosa. Metode deskriptif dilakukan melalui pengujian pada koefisien varian, rasio skewness, rasio kurtosis, histogram, box plot, normal Q-Q plots, dan deterended Q-Q plots. Sedangkan metode analitis dilakukan melalui pengujian nilai Kolmogorov-Smirnov dan nilai Shapiro-Wilk. Suatu data dinyatakan berdistribusi normal apabila memenuhi kriteria yang ditunjukkan pada tabel 2. [22]

Tabel 2 Kriteria data berdistribusi normal

Metode	Parameter	Kriteria	Keterangan
Deskriptif	Koefisien Varian (%)	Nilai koefisien varians <30%.	$\frac{SD}{mean} \times 100\%$
	Rasio Skewness	Nilai rasio skewness -2 s/d 2.	$\frac{Skewness}{SE Skewness}$
	Rasio Kurtosis	Nilai rasio kurtosis -2 s/d 2.	$\frac{Kurtosis}{SE Kurtosis}$
	Histogram	Simetris tidak miring ke kiri maupun ke kanan, tidak terlalu tinggi atau terlalu rendah.	
	Box Plot	Simetris median tepat di tengah, tidak ada outlier atau nilai ekstrim.	
	Normal Q-Q plots	Data menyebar sekitar garis.	
	Deterended Q-Q plots	Data menyebar sekitar garis pada nilai 0.	
Analitik	Kolmogorov-Smirnov	Nilai kemaknaan (p) > 0.05.	Untuk sampel besar (>50)
	Shapiro-Wilk	Nilai kemaknaan (p) > 0.05.	Untuk sampel kecil (≤50)

Keterangan:

SD=Standar Deviasi

SE=Standar Error

2.2.4 Uji Homogenitas

Uji homogenitas adalah pengujian mengenai varian dan digunakan untuk mengetahui apakah kedua kelompok sampel mempunyai varian yang sama atau tidak. Data yang lulus uji homogenitas dapat dianggap bahwa data variabel tersebut pada masing-masing jenis penyakit berasal dari populasi yang

homogen. Uji homogenitas merupakan salah satu prasyarat dalam analisis Independent Sample T Test dan Anova. Syarat pengambilan keputusan dalam uji homogenitas adalah sebagai berikut: [21]

- Jika nilai signifikansi <0.05 , maka dapat dikatakan bahwa varian dari dua atau lebih kelompok populasi data adalah tidak sama.
- Jika nilai signifikansi >0.05 , maka dapat dikatakan bahwa varian dari dua atau lebih kelompok populasi data adalah sama.

Langkah-langkah menghitung uji homogenitas adalah sebagai berikut: [21]

- a. Mencari varians (standar deviasi) pada masing-masing variabel dengan persamaan (6):

$$\begin{aligned} \text{standar deviasi} &= \sqrt{\text{varian}} \\ \text{standar deviasi}^2 &= \text{varian} \\ s^2 &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \text{ atau } s^2 = \frac{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}{n(n-1)} \dots (6) \end{aligned}$$

- b. Menghitung rasio varians pada masing-masing variabel dengan persamaan (7).

$$F = \frac{s_{\text{besar}}}{s_{\text{kecil}}} \dots (7)$$

- c. Membandingkan F hitung dengan tabel distribusi F.
 - Untuk varians dari kelompok dengan variance terbesar adalah dk pembilang n-1.
 - Untuk varians dari kelompok dengan variance terkecil adalah dk penyebut n-1.
 - Jika F hitung $<$ Tabel F, berarti homogen.
 - Jika F hitung $>$ Tabel F, berarti tidak homogen.

2.2.5 Uji Analysis of Variances

Analisis ragam (uji F) pada penelitian ini dilakukan untuk menguji kesamaan rata-rata antar kelompok variabel diagnosa. Dengan menggunakan metode Anova, pengujian kesamaan rata-rata beberapa kelompok dapat dilakukan lebih cepat dan beresiko mengandung kesalahan yang lebih kecil. Perbedaan uji

anova dengan uji-t adalah pada jumlah kelompok yang diuji. Uji-t hanya dapat digunakan untuk menguji maksimal rata-rata 2 (dua) kelompok, sedangkan uji anova dapat menguji rata-rata beberapa kelompok. Uji anova hanya dapat digunakan untuk jenis data parametrik, berdistribusi normal dan bersifat homogen. [18]

Berikut ini merupakan langkah-langkah dalam pengujian anova: [18]

Langkah 1

Menentukan hipotesis nol dan hipotesis alternatifnya.

Langkah 2

Menentukan tingkat signifikansi (0.01 atau 0.05).

Langkah 3

Menentukan aturan pengambilan keputusan dengan nilai kritis F.

Langkah 4

Perhitungan uji anova yang dirangkum dalam tabel 3.

Tabel 3 Rumus uji anova

Sumber variasi	Jumlah kuadrat	Derajat kebebasan	Kuadrat Rata-rata	F
Perlakuan	SST	k-1	$MST = \frac{SST}{k-1}$	$\frac{MST}{MSE}$
Kesalahan	SSE	n-k	$MSE = \frac{SSE}{n-k}$	
Total	SS total	n-1		

Keterangan:

SST=Sum Square for Treatment

SSE=Sum Square Error

k=jumlah kelas

n=jumlah data

MST=Mean Square for Treatment

MSE=Mean Square Error

2.2.6 Uji Korelasi

Uji korelasi merupakan suatu teknik statistic yang digunakan untuk mengetahui keeratan hubungan antar variabel dengan hasil yang sifatnya kuantitatif, serta untuk mengetahui arah hubungan yang terjadi [23]. Uji korelasi terdiri dari beberapa variasi, seperti Pearson dan Spearman's rho. Untuk uji korelasi variabel yang bersifat parametrik menggunakan uji korelasi Pearson, sedangkan untuk uji korelasi variabel yang bersifat non-parametrik menggunakan uji korelasi Spearman. Untuk menghitung nilai korelasi pearson menggunakan persamaan (8), sedangkan untuk menghitung nilai korelasi spearman menggunakan persamaan (10). [24]

- a. Koefisien Korelasi Pearson atau Koefisien Korelasi Pearson Product Moment.

$$r_{xy} = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2} \sqrt{n\sum y^2 - (\sum y)^2}} \dots (8)$$

Rumus uji signifikansi korelasi Pearson Product Moment

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \dots (9)$$

keterangan:

x=variabel bebas

y=variabel terikat

n=jumlah data

- b. Koefisien Korelasi Spearman atau Koefisien Korelasi Range

$$r_{\text{rank}} = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \dots (10)$$

keterangan:

d_i =selisih dari rank variabel x dan y

n=banyaknya pasangan rank

Tabel 4 menunjukkan besaran nilai korelasi yang dihasilkan dari uji korelasi beserta klasifikasinya.

Tabel 4 Besaran nilai korelasi

Nilai Korelasi	Keterangan
$0.90 \leq r < 1.00$ atau $-0.90 \leq r < -1.00$	Sangat kuat
$0.70 \leq r < 0.90$ atau $-0.70 \leq r < -0.90$	Kuat
$0.50 \leq r < 0.70$ atau $-0.50 \leq r < -0.70$	Moderat
$0.30 \leq r < 0.50$ atau $-0.30 \leq r < -0.50$	Lemah
$0.00 \leq r < 0.30$ atau $-0.00 \leq r < -0.30$	Sangat Lemah

Nilai korelasi yang semakin mendekati nilai 1 atau -1 menunjukkan bahwa hubungan antar variabel sangat kuat, sedangkan apabila nilai korelasi mendekati nol (0) berarti hubungan antar variabel sangat lemah. Korelasi bernilai positif menunjukkan hubungan kedua variabel tersebut searah (apabila nilai suatu variabel tinggi maka nilai pada variabel yang lain juga tinggi) sedangkan apabila korelasi bernilai negatif menunjukkan bahwa hubungan kedua variabel tersebut terbalik (apabila nilai suatu variabel tinggi maka nilai pada variabel yang lain rendah).

2.2.7 Penyakit Paru-paru

2.2.7.1 Pengertian Penyakit Paru-paru

Paru-paru merupakan salah satu organ pernapasan yang berfungsi untuk melakukan respirasi, yaitu mengubah gas karbondioksida (CO_2) menjadi gas oksigen (O_2) dan air (H_2O). Sebelum sampai ke paru-paru, udara yang dihirup manusia akan melewati hidung, pangkal laring, kemudian menuju ke kedua bronki utama (bronkus) dan akan disalurkan ke bronki yang paling kecil (bronkioli), selanjutnya udara dimasukkan ke dalam jutaan kantong udara (alveoli) yang berada dalam paru-paru. Paru-paru memiliki lapisan pelindung (pleura) yang juga berfungsi untuk membantu kontraksi dalam rongga dada [7]. Penyakit paru-paru adalah kondisi paru-paru dimana terjadi peradangan atau pengumpulan cairan (darah atau nanah) atau masuknya bakteri, virus atau jamur ke dalam paru-paru yang

kemudian menyebabkan paru-paru tidak berfungsi dengan baik (normal).

2.2.7.2 Jenis Penyakit Paru-paru

Jenis penyakit paru-paru, antara lain:

2.2.7.2.1 Tuberculosis (TB)

Penyakit *tuberculosis* (TB) merupakan penyakit menular yang menyebar melalui udara ketika seseorang dengan infeksi TB aktif batuk, bersin, atau menyebarkan butiran ludah mereka melalui udara. Orang dengan sistem kekebalan tubuh lemah memiliki risiko yang jauh lebih besar untuk terjangkit TB. Penyakit ini disebabkan oleh bakteri *Mycobacterium tuberculosis*. Gejala yang dialami penderita TB paru, antara lain: batuk berdahak lebih dari 2 (dua) minggu, batuk dengan mengeluarkan darah, dada terasa nyeri dan sesak.

2.2.7.2.2 Pneumonia

Radang paru-paru (*pneumonia*) merupakan penyakit pada paru-paru dimana *pulmonary alveolus* (alveoli) yang berfungsi untuk mengubah karbondioksida atau udara yang dihirup manusia menjadi oksigen mengalami peradangan dan terisi cairan. *Pneumonia* disebabkan oleh virus, bakteri, dan jamur. *Pneumonia* dapat dicegah dengan imunisasi, nutrisi, dan faktor lingkungan yang sehat. Gejala yang dialami penderita *pneumonia*, antara lain: batuk berdahak kental dan berwarna kuning, dada terasa nyeri dan sesak, demam tinggi.

2.2.7.2.3 Penyakit Paru Obstruktif Kronis (PPOK)

PPOK merupakan istilah yang digunakan untuk menggambarkan penyakit paru yang sangat buruk, seperti asma jangka panjang (tidak ada perubahan), bronkitis menahun/kronis, emfisema (kondisi kantung udara di paru-paru (alveoli) mengalami kerusakan yang sangat buruk) [6].

Sebagian besar penderita PPOK dikarenakan tembakau dari rokok (baik perokok aktif maupun pasif) dan polusi udara. PPOK tidak dapat disembuhkan, tetapi pengobatan pada penderita dapat meringankan gejala, meningkatkan kualitas hidup dan risiko kematian. [1]

2.2.7.2.4 Bronkitis

Bronkitis merupakan infeksi akibat bakteri atau virus pada saluran pernapasan utama (bronkus), dan biasanya berasal dari hidung dan tenggorokan sehingga menyebabkan terjadinya peradangan pada saluran tersebut. Bronkitis yang berulang (kronis) dapat menjadi salah satu gejala dari PPOK. Gejala yang dialami penderita bronkitis, antara lain: batuk berlendir berwarna abu-abu atau hijau, sakit tenggorokan, dada terasa nyeri dan sesak, mudah lelah, hidung beringus atau tersumbat. [5]

2.2.8 Klasifikasi

Klasifikasi merupakan salah satu teknik dalam *data mining* untuk mengelompokkan data atau pengamatan yang memiliki kesamaan ke target variabel kategori. *Data mining* merupakan proses menemukan hubungan yang berarti, pola dan kecenderungan dari sekumpulan data dengan menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait. [12]

Klasifikasi membutuhkan data berlabel dalam proses pelatihan (training) agar dapat membentuk suatu model algoritma yang dapat digunakan menyelesaikan suatu permasalahan (pengelompokan).

Perbedaan antara klasifikasi dan prediksi terletak pada hasil keluaran (*output*) dari teknik pengelompokannya. Klasifikasi mengelompokkan data ke target variabel kategori yang bersifat diskrit atau real. Sedangkan hasil keluaran dari prediksi berupa nilai-nilai kontinyu atau real yang belum dapat diketahui sebelumnya secara pasti.

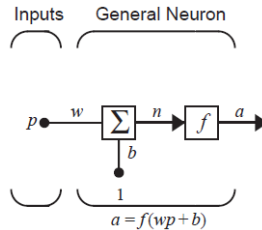
2.2.9 Artificial Neural Network

Artificial Neural Network (ANN) merupakan salah satu sistem kecerdasan buatan yang dibuat untuk membantu manusia dalam menyelesaikan permasalahan yang cukup kompleks, seperti aplikasi pengenalan pola (*pattern recognition*), pembentukan pola (*pattern generation*) dan klasifikasi pola (*pattern classification*) [3].

ANN mempelajari pola-pola dan hubungan data yang jumlahnya sangat besar dan bersifat non-linier dengan menyaring data, mencari hubungan, membangun pola dan mengoreksi kesalahan pola itu sendiri berkali-kali. ANN dikembangkan tidak untuk menyelesaikan kasus tertentu, para pengembang ANN lebih bertujuan untuk menyimpan kecerdasan tersebut ke dalam perangkat keras agar dapat dilatih ulang untuk menyelesaikan banyak permasalahan serupa atau berbeda lainnya. Hal ini berbeda dengan sistem para pakar dalam memecahkan suatu permasalahan, mereka berusaha menyelesaikan kasus tertentu tetapi kecerdasan yang telah dikembangkan tidak dapat dilatih ulang atau dikembangkan untuk menyelesaikan kasus serupa atau berbeda lainnya. [3]

2.2.9.1 Komponen ANN

Bentuk kecerdasan buatan yang dikembangkan meniru cara kerja sistem syaraf biologis manusia (neuron). Dimana komponen dalam setiap sistem ANN dapat dianalogikan pada setiap bagian komponen sistem neuron manusia, seperti dendrit (bagian yang menerima masukan/*input/signal*), badan sel (mengolah masukan), dan akson (mentransmisikan masukan tersebut ke neuron yang lain/*output*). [4]



Gambar 1 Struktur ANN [10]

Setiap informasi yang diterima oleh dendrit (input) akan dijumlahkan dan dikirim melalui akson menuju dendrit (output) untuk dikirimkan ke dendrit (*input*) neuron lainnya. Informasi akan diterima oleh neuron lainnya apabila memenuhi batasan tertentu, yang disebut nilai ambang (*threshold*). Neuron yang telah menerima informasi dan mengirimkan informasi ke neuron lain dapat dikatakan teraktivasi. Neuron akan menerima informasi dari neuron lain dalam bentuk nilai yang disebut bobot. Bobot juga dapat menunjukkan seberapa besar hubungan antar neuron.

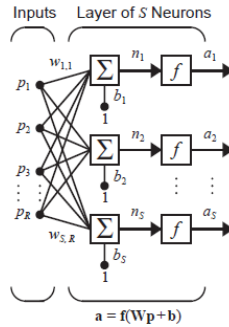
2.2.9.2 Arsitektur ANN

Arsitektur ANN merupakan pola hubungan antar neuron. Neuron yang memiliki pola bobot dan fungsi aktivasi yang sama akan dikelompokkan dalam lapisan (*layer*) yang sama. Biasanya informasi yang dirambatkan dari lapisan input ke lapisan output akan melewati lapisan tersembunyi (*hidden layer*).

Ada beberapa arsitektur jaringan syaraf, antara lain:

2.2.9.2.1 Jaringan dengan lapisan tunggal (*single layer network*)

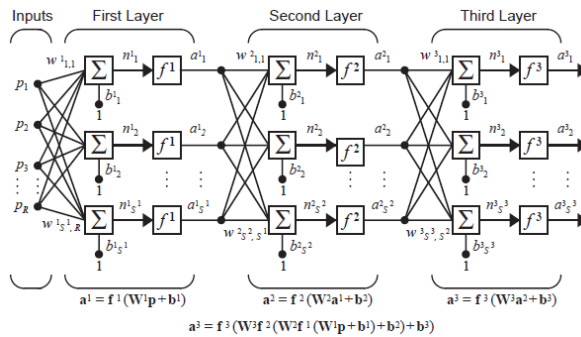
Jaringan dengan lapisan tunggal hanya memiliki satu lapisan yang memuat bobot-bobot yang saling terhubung. Jaringan ini tidak memiliki *hidden layer*, semua unit *input* akan dihubungkan dengan setiap unit *output*.



Gambar 2 Single layer network [10]

2.2.9.2.2 Jaringan dengan banyak lapisan (multilayer network)

Jaringan dengan banyak lapisan memiliki 1 atau lebih lapisan (*hidden layer* dan bobot) di antara lapisan input dan lapisan output. Jaringan dengan banyak lapisan dapat menyelesaikan permasalahan yang lebih kompleks daripada jaringan dengan lapisan tunggal.



Gambar 3 Three-layer network [10]

2.2.10 Recall, Precision dan Accuracy

Recall, *precision* dan *accuracy* merupakan ukuran untuk mengetahui ketepatan (performa) prediksi dari suatu pengklasifikasi dengan lebih dari 2 (dua) kelas target (multikelas). Pada klasifikasi multikelas, akurasi hanya

mengukur ketepatan pengklasifikasi pada seluruh kelas yang ada. Sedangkan, *precision* dapat mengukur ketepatan prediksi pengklasifikasi pada kelas tertentu, dan *recall* dapat mengukur berapa banyak instans dari kelas tertentu yang dapat diprediksi secara benar.

Perhitungan nilai *recall*, *precision* dan *accuracy* dapat dihitung berdasarkan rumus 11-13 [11], dengan bantuan tabel kontingensi (*contingency table*) atau *confusion matrix*.

$$\text{Recall (\%)} = \frac{\text{jumlah positif benar pada kelas tertentu}}{\text{jumlah total data aktual pada kelas tertentu}} \times 100\% \dots (11)$$

$$\text{Precision (\%)} = \frac{\text{jumlah positif benar pada kelas tertentu}}{\text{jumlah total data prediksi pada kelas tertentu}} \times 100\% \dots (12)$$

$$\text{Accuracy (\%)} = \frac{\text{jumlah positif benar} + \text{jumlah negatif benar}}{\text{jumlah total data}} \times 100\% \dots (13)$$

BAB III

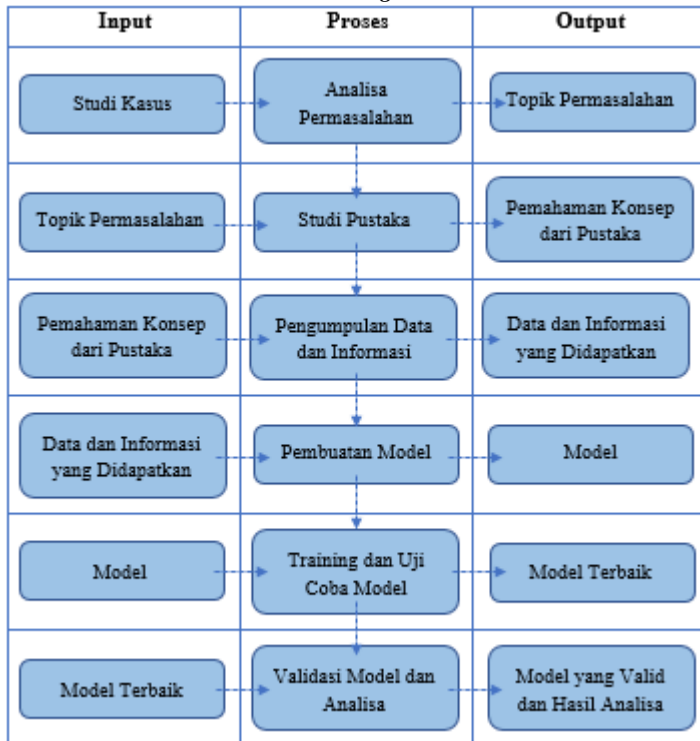
METODOLOGI PENELITIAN

Pada bab ini, akan menjelaskan mengenai metodologi yang yang digunakan sebagai panduan untuk menyelesaikan penelitian tugas akhir ini.

3.1 Diagram Metodologi

Tabel 5 merupakan diagram metodologi penelitian tugas akhir ini.

Tabel 5 Metodologi Penelitian



3.2 Uraian Metodologi

3.2.1 Analisa Permasalahan

Pada tahap ini dilakukan penggalian dan analisa permasalahan yang terdapat pada studi kasus. Adapun permasalahan yang terdapat pada studi kasus RSUD Kertosono adalah belum adanya sistem diagnosa penyakit paru-paru untuk tenaga medis atau pasien secara real time di ruang poli paru sehingga pasien atau tenaga medis harus menunggu hasil uji laboratorium. Hasil uji laboratorium di RSUD Kertosono dapat diketahui sehari setelah uji laboratorium dilakukan karena banyaknya variabel diagnose yang dibutuhkan (rontgen organ paru-paru, darah, dahak). Selain uji laboratorium, tenaga medis juga dapat mendiagnosa jenis penyakit paru pasien dari beberapa variabel keluhan yang dirasakan secara langsung. Oleh karena itu, penelitian ini diharapkan dapat membantu tenaga medis dan pasien untuk lebih akurat dan lebih cepat dalam mendiagnosa jenis penyakit paru-paru yang diderita pasien karena variabel yang digunakan untuk mendiagnosa tidak sebanyak yang dibutuhkan dalam uji laboratorium sehingga pasien dapat lebih cepat diobati atau ditangani.

3.2.2 Studi Pustaka

Studi pustaka (*Library Research*) ini dilakukan dari berbagai referensi, seperti buku pustaka, makalah, jurnal ilmiah, paper pada penelitian sebelumnya, dan narasumber, serta beberapa referensi lainnya terkait penyakit paru-paru dan metode *Artificial Neural Network* (ANN) dalam mendiagnosa penyakit (khususnya paru-paru). Studi pustaka ini dilakukan untuk meningkatkan pemahaman dan teori dasar yang mendukung pencapaian tujuan penelitian tugas akhir ini.

3.2.3 Pengumpulan Data dan Informasi

Untuk membuat model diagnosa jenis penyakit paru-paru dengan metode *Artificial Neural Network* (ANN), maka dibutuhkan data dan informasi terkait rekam medis sejumlah pasien penderita penyakit paru-paru. Penelitian ini

mendapatkan data sampel dari poli paru RSUD Kertosono sejumlah 300 data rekam medis dari kurun waktu Januari 2014 hingga Desember 2016.

3.2.4 Pembuatan Model

Data rekam medis yang didapat yaitu sejumlah 7 input (jenis kelamin, umur, berat badan, systole, diastole, batuk, sesak) dan 4 output (*tuberculosis*, PPOK, *pneumonia*, bronkitis).

3.2.5 Training dan Uji Coba Model

Setelah model dibuat, tahapan selanjutnya yaitu melatih dan menguji model tersebut hingga tercapai model terbaik. Untuk menguji model tersebut, data sampel akan dibagi menjadi 2 (dua) bagian, yaitu data training dan data testing dengan masing-masing proporsinya yaitu 75% data sampel untuk data training dan 25% data sampel untuk data testing.

3.2.6 Validasi Model dan Analisa

Setelah model dilatih dan diuji dengan data sampel, hasilnya akan dibandingkan dengan menilai tingkat *precision*, *recall* dan *accuracy* [11]. Model dengan tingkat *precision*, *recall* dan *accuracy* paling tinggi akan direkomendasikan sebagai dasar untuk pembuatan sistem diagnosa penyakit paru-paru. Analisa dilakukan untuk penarikan kesimpulan dari hasil validitas model sehingga dapat dijadikan saran untuk mengembangkan model agar lebih baik lagi.

Halaman ini sengaja dikosongkan

BAB IV

PERANCANGAN

4.1 Persiapan Data

Data yang digunakan pada penelitian ini diperoleh dari data rekam medis penderita penyakit paru yang sedang menjalani pengobatan rawat jalan di RSUD Kertosono selama kurun waktu 2014-2016. Jumlah data yang digunakan dalam penelitian ini sebanyak 483 data, dengan pembagian jumlah data yang digunakan sebagai pelatihan model ANN maupun uji coba model ANN adalah sebagai berikut.

Tabel 6 Pembagian Data Training dan Testing

Jenis Penyakit	Jumlah Data	Training (75%)	Testing (25%)
PPOK	72	54	18
Tuberculosis	316	237	79
Bronkitis	92	69	23
Pneumonia	3	2	1
Jumlah	483	362	121

Data rekam medis penderita penyakit paru terdiri dari 18 (delapan belas) variabel, yaitu jenis kelamin, umur, berat badan, sistole, diastole, batuk, sesak, demam, mual, pusing, nyeri dada, nyeri perut, nyeri kepala, nyeri kaki (bengkak), nyeri tenggorokan, nyeri punggung, pengaburan mata, dan nyeri leher (bengkak). Data rekam medis pasien penderita penyakit paru yang dimiliki RSUD Kertosono tercatat ke dalam 2 (dua) bentuk penyimpanan, yaitu dalam bentuk formulir (manual) dan database (digital). Namun, tidak semua variabel rekam medis yang tercatat di formulir disimpan dalam bentuk digital. Di dalam database (digital) hanya mencatat riwayat penyakit yang dialami pasien, tidak mencatat gejala (variabel diagnosa) yang dialami pasien. Selain itu, data diagnosa yang tercatat dalam formulir rekam medis pasien tidak semuanya terisi secara lengkap. Hal tersebut disebabkan oleh beberapa kemungkinan, seperti kelalaian dokter (tenaga medis), pasien benar-benar

tidak mengalami gejala terkait atau kemungkinan yang lain. Oleh sebab itu, data rekam medis yang diperoleh terdapat nilai yang hilang atau kosong (*missing value*).

Dalam proses data mining, data yang digunakan hendaknya tidak mengandung data yang *error*, tidak lengkap (*missing value*), *not available* (kosong). Pada penelitian ini, metode yang digunakan untuk menangani nilai yang kosong adalah interpolasi. Metode interpolasi ini dipilih karena dalam suatu set data variabel antar diagnose dianggap memiliki hubungan (korelasi), dimana hubungan tersebut akan menghasilkan suatu grafik (fungsi) yang berbeda antar jenis penyakit paru. Sehingga dari grafik (fungsi) tersebut dapat diperkirakan nilai yang paling optimal untuk mengisi data yang kosong. Jenis interpolasi yang digunakan adalah jenis interpolasi sederhana (orde 1) atau interpolasi linier.

Dalam melakukan interpolasi linier diperlukan pasangan variabel yang memiliki nilai korelasi paling tinggi agar fungsi yang dihasilkan memiliki keakuratan yang maksimal. Tabel 7 memuat pasangan variabel yang digunakan dalam perhitungan interpolasi. Variabel 1 menunjukkan variabel diagnosa yang memuat nilai kosong (*missing value*) sedangkan variabel 2 menunjukkan pasangan variabel yang memiliki nilai korelasi paling tinggi terhadap variabel 1 seperti yang ditunjukkan pada tabel 7. Selain itu, pada tabel 7 juga memuat keterangan hubungan (korelasi) yang dimiliki oleh pasangan variabel tersebut.

Tabel 7 Nilai korelasi antar variabel diagnosa

Variabel 1	Variabel 2	Nilai Korelasi	Keterangan
Umur	Berat Badan	-0.264	Lemah
Berat Badan	Umur	-0.264	Lemah
Sistole	Diastole	0.701	Cukup Kuat
Diastole	Sistole	0.701	Cukup Kuat
Batuk	Sesak	0.067	Sangat Lemah
Sesak	Umur	0.140	Sangat Lemah
Demam	Mual	0.012	Sangat Lemah

Variabel 1	Variabel 2	Nilai Korelasi	Keterangan
Mual	Nyeri Perut	0.605	Cukup Kuat
Pusing	Nyeri Kepala	0.710	Cukup Kuat
Nyeri Dada	Sesak	0.767	Cukup Kuat
Nyeri Perut	Mual	0.605	Cukup Kuat
Nyeri Kepala	Pusing	0.710	Cukup Kuat
Nyeri Kaki	Nyeri Punggung	0.012	Sangat Lemah
Nyeri Tenggorokan	Batuk	0.603	Cukup Kuat
Nyeri Punggung	Nyeri Kaki	0.067	Sangat Lemah
Pengaburan Mata	Pusing	0.304	Lemah
Nyeri Leher	Batuk	0.600	Cukup Kuat

Berdasarkan hasil interpolasi, nilai pada variabel demam, mual, pusing, nyeri dada, nyeri perut, nyeri kepala, nyeri kaki (bengkak), nyeri tenggorokan, nyeri punggung, pengaburan mata, dan nyeri leher (bengkak) tidak memiliki ragam (variance) yang signifikan. Oleh sebab itu, variabel tersebut tidak digunakan dalam proses selanjutnya (uji hipotesa statistika dan pembuatan model).

Tahapan selanjutnya setelah data yang digunakan dalam penelitian ini tidak memuat nilai yang kosong adalah tahapan pengujian hipotesa statistika (uji beda dan uji asosiasi). Uji hipotesa statistika diperlukan untuk mengenali karakteristik masing-masing variabel dari data set penelitian. Uji hipotesa statistika yang dilakukan antara lain uji normalitas, uji homogenitas, uji anova dan uji korelasi. Berikut ini merupakan hasil dari uji hipotesa statistika yang telah dilakukan.

4.1.1 Uji Normalitas

Uji normalitas pada penelitian ini dilakukan untuk mengetahui persebaran satu set data pada masing-masing variabel, apakah berdistribusi normal atau tidak. Apabila data dinyatakan berdistribusi normal, data tersebut dapat dianggap mewakili populasi. Uji normalitas dilakukan dengan metode deskriptif dan analitis pada setiap variabel diagnosa. Metode deskriptif dilakukan melalui pengujian pada koefisien varian, rasio skewness, rasio kurtosis, histogram, box plot, normal Q-Q plots,

dan deterrended Q-Q plots. Sedangkan metode analitis dilakukan melalui pengujian nilai Kolmogorov-Smirnov dan nilai Shapiro-Wilk. Dalam studi kasus penelitian ini, apabila data memenuhi minimal 1 (satu) parameter dari 9 (sembilan) parameter yang digunakan maka data tersebut dapat disimpulkan berdistribusi normal.

Variabel Jenis Kelamin

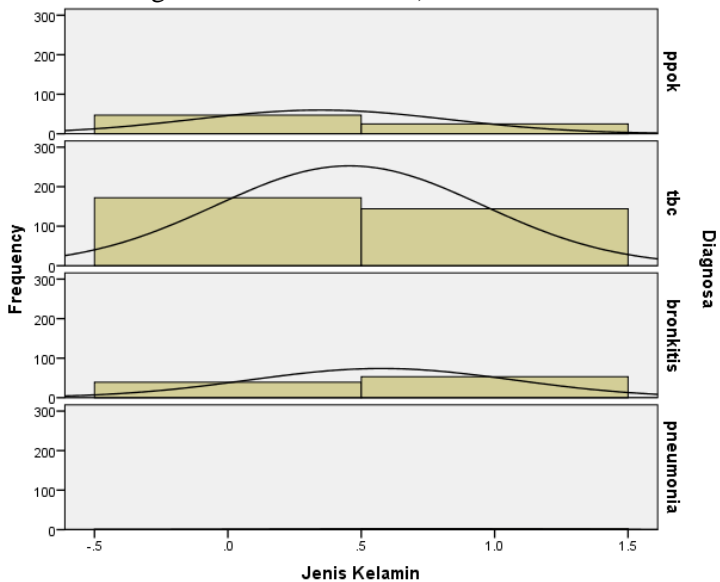
Berdasarkan hasil uji normalitas yang ditunjukkan pada tabel 8, dapat disimpulkan bahwa data pada variabel jenis kelamin pada setiap kelas penyakit paru memiliki persebaran data yang normal. Karena nilai Kolmogorov-Smirnov dan nilai Shapiro-Wilk pada masing-masing kelas penyakit bernilai >0.05 . Selain itu, persentase koefisien varian pada jenis penyakit PPOK, tuberculosis dan bronkitis $<30\%$ dan nilai rasio skewness pada variabel jenis kelamin tuberculosis, bronkitis, pneumonia berada dalam interval -2 dan 2. Hal tersebut membuktikan bahwa data pada variabel jenis kelamin pada setiap kelas penyakit paru dapat dianggap mewakili populasi.

Tabel 8 Hasil uji normalitas variabel jenis kelamin

Parameter	PPOK	Tuberculosis	Bronkitis	Pneumonia
Koefisien Varian (%)	16.28*	6.08*	8.96*	49.7
Rasio Skewness	2.31	1.30*	-1.25*	1.41*
Rasio Kurtosis	2.89	-7.26	-3.90	-
Histogram	Simetris*	Simetris*	Simetris*	Simetris*
Box Plot	Tidak simetris	Tidak simetris	Tidak simetris	Tidak simetris
Normal Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Deterrended Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Kolmogoro v-Smirnov	0.418*	0.364*	0.379*	0.385*
Shapiro-Wilk	0.602*	0.634*	0.628*	0.75*
Keterangan	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal

*. Data berdistribusi normal

Histogram pada gambar 4 menggambarkan persebaran data pada variabel jenis kelamin di setiap jenis penyakit paru. Histogram tersebut merupakan salah satu parameter yang digunakan untuk uji normalitas. Apabila histogram simetris (tidak condong ke kanan atau ke kiri) maka dapat disimpulkan bahwa variabel jenis kelamin memiliki distribusi data yang normal. Dari gambar tersebut dapat disimpulkan bahwa persebaran data pada masing-masing jenis penyakit berdistribusi normal karena grafik yang dihasilkan simetris (tidak condong ke kanan atau ke kiri).



Gambar 4 Histogram persebaran data variabel jenis kelamin

Variabel Umur

Berdasarkan hasil uji normalitas yang ditunjukkan pada tabel 9, dapat disimpulkan bahwa data pada variabel umur pada setiap kelas penyakit paru memiliki persebaran data yang normal. Karena nilai Shapiro-Wilk pada masing-masing kelas penyakit bernilai >0.05 dan nilai Kolmogorov-Smirnov pada masing-masing kelas penyakit paru bernilai >0.05 , kecuali penyakit

tuberculosis. Selain itu, persentase koefisien varian pada setiap jenis penyakit bernilai $<30\%$ dan nilai rasio skewness pada variabel umur tuberculosis, bronkitis, pneumonia berada dalam interval -2 dan 2. Hal tersebut membuktikan bahwa data pada variabel umur pada setiap kelas penyakit paru dapat dianggap mewakili populasi.

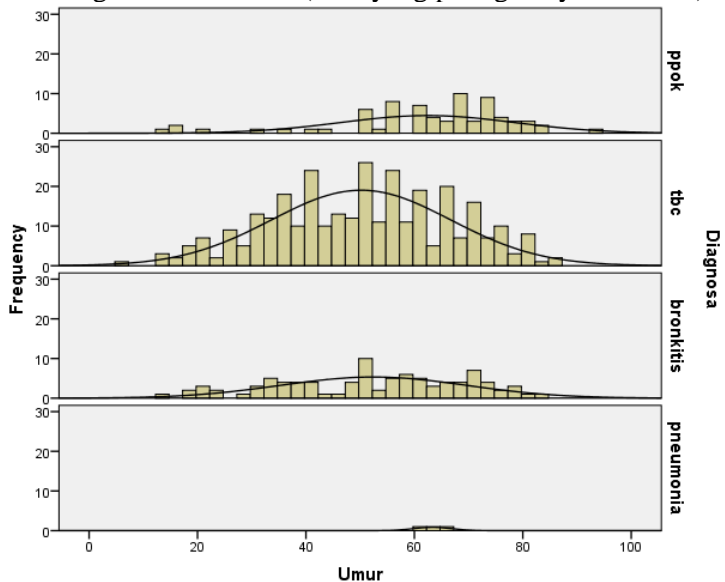
Tabel 9 Hasil uji normalitas variabel umur

Parameter	PPOK	Tuberculosis	Bronkitis	Pneumonia
Koefisien Varian (%)	3.03*	1.85*	3.41*	3.20*
Rasio Skewness	4.551	-0.693*	-1.239*	0.345*
Rasio Kurtosis	3.553	-2.380	-1.524*	-
Histogram	Condong ke kanan	Simetris*	Simetris*	Simetris*
Box Plot	Terdapat outlier di bawah box	Simetris*	Simetris*	Simetris*
Normal Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Deterrended Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Kolmogorov-Smirnov	0.122*	0.049	0.070*	0.204*
Shapiro-Wilk	0.897*	0.989*	0.970*	0.993*
Keterangan	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal

*. Data berdistribusi normal

Histogram pada gambar 5 menggambarkan persebaran data pada variabel umur di setiap jenis penyakit paru. Histogram tersebut merupakan salah satu parameter yang digunakan untuk uji normalitas. Apabila histogram simetris (tidak condong ke kanan atau ke kiri) maka dapat disimpulkan bahwa variabel umur memiliki distribusi data yang normal. Dari histogram tersebut dapat disimpulkan bahwa persebaran data pada jenis penyakit tuberculosis, bronkitis dan pneumonia berdistribusi normal karena grafik yang dihasilkan simetris (tidak condong ke kanan atau ke kiri). Persebaran data pada jenis penyakit ppok lebih condong ke kanan, hal ini menunjukkan bahwa nilai rata-

rata pada variabel umur pada jenis penyakit ppok lebih kecil dibandingkan nilai modus (nilai yang paling banyak muncul).



Gambar 5 Histogram persebaran data variabel umur

Variabel Berat Badan

Berdasarkan hasil uji normalitas yang ditunjukkan pada tabel 10, dapat disimpulkan bahwa data pada variabel berat badan pada setiap kelas penyakit kecuali pneumonia memiliki persebaran data yang normal. Karena nilai Shapiro-Wilk pada kelas penyakit PPOK, tuberculosis, pneumonia dan bronkitis bernilai >0.05 . Selain itu, nilai Kolmogorov-Smirnov pada jenis penyakit bronkitis dan pneumonia bernilai >0.05 . Persentase koefisien varian pada jenis penyakit PPOK, tuberculosis, pneumonia dan bronkitis $<30\%$. Nilai rasio skewness pada tuberculosis, pneumonia dan bronkitis berada dalam interval -2 dan 2. Nilai rasio kurtosis pada penyakit PPOK, tuberculosis dan bronkitis juga berada dalam interval -2 dan 2. Hal tersebut membuktikan bahwa data pada variabel jenis kelamin pada

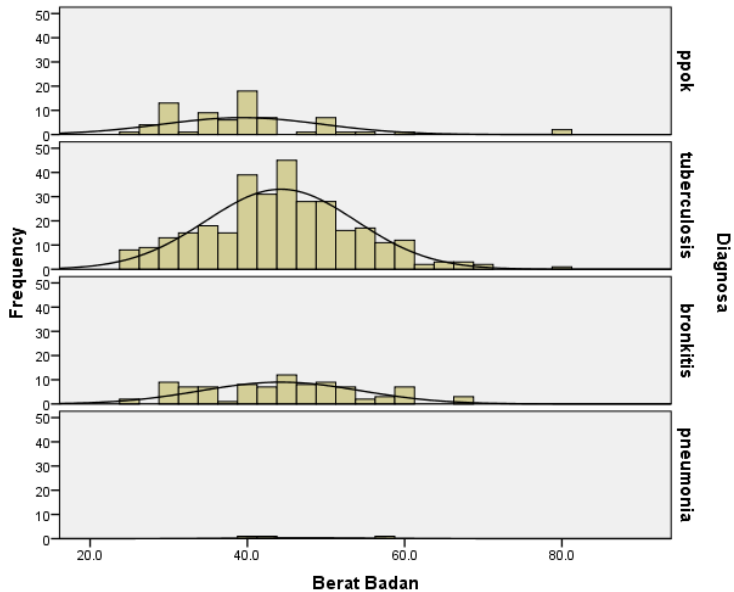
kelas penyakit PPOK, tuberculosis, pneumonia dan bronkitis dapat dianggap mewakili populasi.

Tabel 10 Hasil uji normalitas variabel berat badan

Parameter	PPOK	Tuberculosis	Bronkitis	Pneumonia
Koefisien Varian (%)	1.19*	1.16*	5.41*	11.33*
Rasio Skewness	2.70	1.42*	0.74*	1.34*
Rasio Kurtosis	0.93*	-0.20*	-0.83*	-
Histogram	Simetris*	Simetris*	Simetris*	Simetris*
Box Plot	Terdapat outlier di atas box	Terdapat outlier di atas box	Tidak simetris	Tidak simetris
Normal Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Deterended Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Kolmogorov-Smirnov	0.071	0.031	0.100*	0.346*
Shapiro-Wilk	0.956*	0.993*	0.960*	0.837*
Keterangan	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal

*. Data berdistribusi normal

Histogram pada gambar 6 menggambarkan persebaran data pada variabel berat badan di setiap jenis penyakit paru. Histogram tersebut merupakan salah satu parameter yang digunakan untuk uji normalitas. Apabila histogram simetris (tidak condong ke kanan atau ke kiri) maka dapat disimpulkan bahwa variabel berat badan memiliki distribusi data yang normal. Dari gambar tersebut dapat disimpulkan bahwa persebaran data pada penyakit ppok, tuberculosis, pneumonia dan bronkitis berdistribusi normal karena grafik yang dihasilkan simetris (tidak condong ke kanan atau ke kiri).



Gambar 6 Histogram persebaran data variabel berat badan

Variabel Sistole

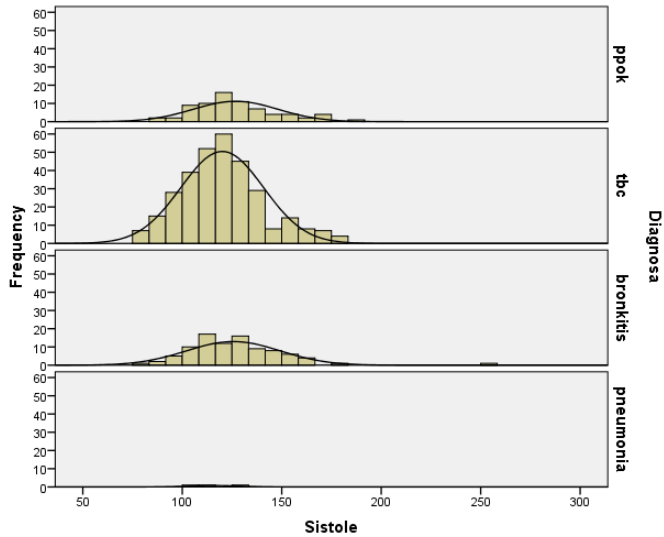
Berdasarkan hasil uji normalitas yang ditunjukkan pada tabel 11, dapat disimpulkan bahwa data pada variabel sistole pada setiap kelas penyakit paru memiliki persebaran data yang normal. Karena nilai Kolmogorov-Smirnov dan nilai Shapiro-Wilk pada masing-masing kelas penyakit bernilai >0.05 . Selain itu, persentase koefisien varian pada setiap jenis penyakit $<30\%$, nilai rasio kurtosis pada variabel ppok, tuberculosis dan bronkitis berada dalam interval -2 dan 2, serta nilai rasio skewness pada variabel pneumonia berada dalam interval -2 dan 2. Hal tersebut membuktikan bahwa data pada variabel sistole pada setiap kelas penyakit paru dapat dianggap mewakili populasi.

Tabel 11 Hasil uji normalitas variabel sistole

Parameter	PPOK	Tuberculosis	Bronkitis	Pneumonia
Koefisien Varian (%)	1.99*	0.97*	1.96*	6.50*
Rasio Skewness	2.41	3.65	6.40	1.33*
Rasio Kurtosis	0.45*	0.44*	14.56	-
Histogram	Simetris*	Simetris*	Simetris*	Simetris*
Box Plot	Terdapat outlier di atas box	Terdapat outlier di atas box	Terdapat nilai ekstrem di atas box	Tidak simetris
Normal Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Deterended Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Kolmogorov-Smirnov	0.124*	0.076*	0.095*	0.343*
Shapiro-Wilk	0.961*	0.977*	0.901*	0.842*
Keterangan	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal

*. Data berdistribusi normal

Histogram pada gambar 7 menggambarkan persebaran data pada variabel sistole di setiap jenis penyakit paru. Histogram tersebut merupakan salah satu parameter yang digunakan untuk uji normalitas. Apabila histogram simetris (tidak condong ke kanan atau ke kiri) maka dapat disimpulkan bahwa variabel sistole memiliki distribusi data yang normal. Dari gambar tersebut dapat disimpulkan bahwa persebaran data pada masing-masing jenis penyakit berdistribusi normal karena grafik yang dihasilkan simetris (tidak condong ke kanan atau ke kiri).



Gambar 7 Histogram persebaran data variabel sistole

Variabel Diastole

Berdasarkan hasil uji normalitas yang ditunjukkan pada tabel 12, dapat disimpulkan bahwa data pada variabel jenis kelamin pada setiap kelas penyakit paru memiliki persebaran data yang normal. Karena nilai Kolmogorov-Smirnov dan nilai Shapiro-Wilk pada masing-masing kelas penyakit bernilai >0.05 . Selain itu, persentase koefisien varian pada jenis penyakit PPOK, tuberculosis dan bronkitis $<30\%$ dan nilai rasio skewness pada jenis kelamin tuberculosis, bronkitis, pneumonia berada dalam interval -2 dan 2. Hal tersebut membuktikan bahwa data pada variabel jenis kelamin pada setiap kelas penyakit paru dapat dianggap mewakili populasi.

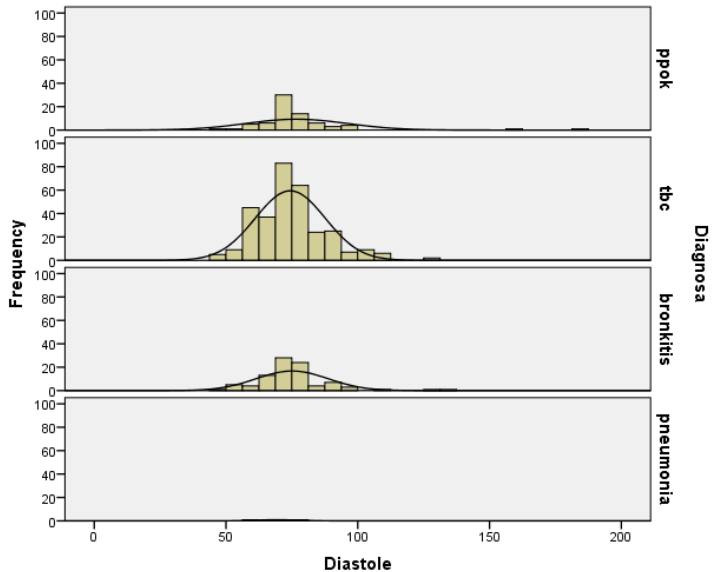
Tabel 12 Hasil uji normalitas variabel diastole

Parameter	PPOK	Tuberculosis	Bronkitis	Pneumonia
Koefisien Varian (%)	2.96	1.00	1.90	6.67
Rasio Skewness	12.75	6.45	6.05	-0.30

Parameter	PPOK	Tuberculosis	Bronkitis	Pneumonia
Rasio Kurtosis	32.17	5.99	10.48	-
Histogram	Condong ke kiri	Condong ke kiri	Condong ke kiri	simetris
Box Plot	Terdapat nilai ekstrem di atas box dan outlier di bawah box	Terdapat outlier dan nilai ekstrem di atas box	Terdapat nilai ekstrem di atas box dan outlier di atas dan di bawah box	simetris
Normal Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Deterended Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Kolmogorov-Smirnov	0.204	0.126	0.154	0.2
Shapiro-Wilk	0.652	0.954	0.884	0.995
Keterangan	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal

*. Data berdistribusi normal

Histogram pada gambar 8 menggambarkan persebaran data pada variabel diastole di setiap jenis penyakit paru. Histogram tersebut merupakan salah satu parameter yang digunakan untuk uji normalitas. Apabila histogram simetris (tidak condong ke kanan atau ke kiri) maka dapat disimpulkan bahwa variabel diastole memiliki distribusi data yang normal. Dari gambar tersebut dapat disimpulkan bahwa persebaran data pada masing-masing jenis penyakit berdistribusi normal karena grafik yang dihasilkan simetris (tidak condong ke kanan atau ke kiri).



Gambar 8 Histogram persebaran data variabel diastole

Variabel Batuk

Berdasarkan hasil uji normalitas yang ditunjukkan pada tabel 13, dapat disimpulkan bahwa data pada variabel batuk pada setiap kelas penyakit paru memiliki persebaran data yang normal. Karena nilai Kolmogorov-Smirnov dan nilai Shapiro-Wilk pada masing-masing kelas penyakit bernilai >0.05 . Selain itu, persentase koefisien varian pada jenis penyakit PPOK, tuberculosis dan bronkitis $<30\%$ dan nilai rasio skewness variabel batuk pada jenis kelamin bronkitis dan pneumonia berada dalam interval -2 dan 2 . Hal tersebut membuktikan bahwa data pada variabel batuk pada setiap kelas penyakit paru dapat dianggap mewakili populasi.

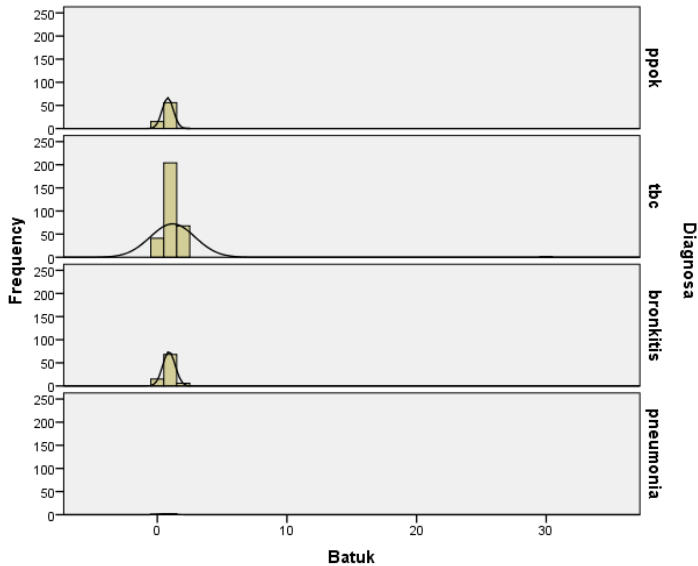
Tabel 13 Hasil uji normalitas variabel batuk

Parameter	PPOK	Tuberculosis	Bronkitis	Pneumonia
Koefisien Varian (%)	6.29*	8.30*	5.67*	49.7

Parameter	PPOK	Tuberculosis	Bronkitis	Pneumonia
Rasio Skewness	-3.63	107.04	-1.18*	-1.41*
Rasio Kurtosis	1.01	898.43	2.45	-
Histogram	Simetris*	Simetris*	Simetris*	Simetris*
Box Plot	Terdapat nilai ekstrem di atas dan di bawah box	Terdapat nilai ekstrem di atas dan di bawah box	Terdapat nilai ekstrem di atas dan di bawah box	Tidak simetris
Normal Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Deterended Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Kolmogorov-Smirnov	0.465*	0.328*	0.415*	0.385*
Shapiro-Wilk	0.566*	0.206*	0.647*	0.750*
Keterangan	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal

*. Data berdistribusi normal

Histogram pada gambar 8 menggambarkan persebaran data pada variabel batuk di setiap jenis penyakit paru. Histogram tersebut merupakan salah satu parameter yang digunakan untuk uji normalitas. Apabila histogram simetris (tidak condong ke kanan atau ke kiri) maka dapat disimpulkan bahwa variabel batuk memiliki distribusi data yang normal. Dari gambar tersebut dapat disimpulkan bahwa persebaran data pada masing-masing jenis penyakit berdistribusi normal karena grafik yang dihasilkan simetris (tidak condong ke kanan atau ke kiri).



Gambar 9 Histogram persebaran data variabel batuk

Variabel Sesak

Berdasarkan hasil uji normalitas yang ditunjukkan pada tabel 14, dapat disimpulkan bahwa data pada variabel jenis kelamin pada setiap kelas penyakit paru memiliki persebaran data yang normal. Karena nilai Kolmogorov-Smirnov dan nilai Shapiro-Wilk pada masing-masing kelas penyakit bernilai >0.05 . Selain itu, persentase koefisien varian pada jenis penyakit PPOK, tuberculosis dan bronkitis $<30\%$ dan nilai rasio skewness variabel sesak pada jenis kelamin bronkitis dan pneumonia berada dalam interval -2 dan 2 . Hal tersebut membuktikan bahwa data pada variabel sesak pada setiap kelas penyakit paru dapat dianggap mewakili populasi.

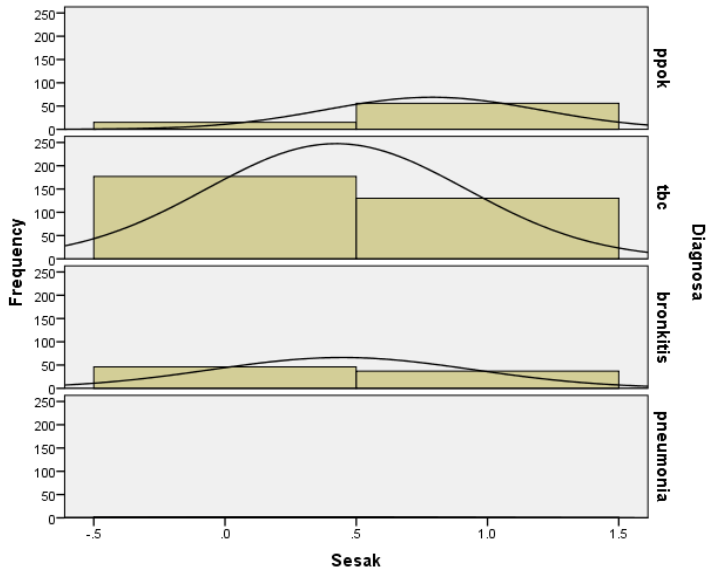
Tabel 14 Hasil uji normalitas variabel sesak

Parameter	PPOK	Tuberculosis	Bronkitis	Pneumonia
Koefisien Varian (%)	6.20*	6.67*	12.22*	100

Parameter	PPOK	Tuberculosis	Bronkitis	Pneumonia
Rasio Skewness	-5.07	2.37	0.84*	1.41*
Rasio Kurtosis	0.16*	-6.91	-3.82	-
Histogram	Condong ke kanan	Condong ke kiri	Condong ke kiri	Condong ke kiri
Box Plot	Terdapat nilai ekstrem di bawah box	Tidak simetris	Tidak simetris	Tidak simetris
Normal Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Deterended Q-Q plots	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis	Terdapat data di luar garis
Kolmogorov-Smirnov	0.485*	0.380*	0.368*	0.385*
Shapiro-Wilk	0.502*	0.628*	0.632*	0.750*
Keterangan	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal	Berdistribusi normal

*. Data berdistribusi normal

Histogram pada gambar 10 menggambarkan persebaran data pada variabel sesak di setiap jenis penyakit paru. Histogram tersebut merupakan salah satu parameter yang digunakan untuk uji normalitas. Apabila histogram simetris (tidak condong ke kanan atau ke kiri) maka dapat disimpulkan bahwa variabel sesak memiliki distribusi data yang normal. Dari gambar tersebut dapat disimpulkan bahwa persebaran data pada masing-masing jenis penyakit berdistribusi normal karena grafik yang dihasilkan simetris (tidak condong ke kanan atau ke kiri).



Gambar 10 Histogram persebaran data variabel sesak

Berdasarkan hasil uji normalitas, dapat disimpulkan bahwa variabel jenis kelamin, umur, berat badan, sistole, diastole, batuk, sesak memiliki persebaran data yang normal (berdistribusi normal).

Tabel 15 Kesimpulan hasil uji normalitas

Jenis Penyakit	Normalitas
Jenis Kelamin	✓
Umur	✓
Berat Badan	✓
Sistole	✓
Diastole	✓
Batuk	✓
Sesak	✓

4.1.2 Uji Homogenitas

Uji homogenitas (uji Levene's) varian pada penelitian ini dilakukan untuk mengetahui kesamaan atau ketidaksamaan varians (ragam) suatu set data, apabila data pada masing-masing

variabel diagnosa memiliki varian yang homogen maka akan dilakukan pengujian selanjutnya, salah satunya yaitu Anova (uji F). Syarat suatu data disebut homogen apabila nilai signifikansi $F > 0.05$. Data yang lulus uji homogenitas dapat dianggap bahwa data variabel tersebut pada masing-masing jenis penyakit berasal dari populasi yang homogen. Data yang dapat dilakukan uji homogenitas adalah data yang bersifat parametrik. Pada studi kasus penelitian ini, variabel diagnosa yang bersifat parametrik adalah variabel umur, berat badan, sistole, diastole.

Variabel Umur

Nilai varians (standar deviasi) variabel umur pada setiap jenis penyakit ditunjukkan pada tabel 16. Semakin tinggi nilai standar deviasi pada masing jenis penyakit, menunjukkan semakin tinggi pula perbedaan rata-rata pada masing-masing jenis penyakit.

Tabel 16 Deskripsi statistik data variabel umur

					95% Confidence Interval for Mean		
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum Maximum
ppok	72	62.42	16.101	1.897	58.63	66.20	13 93
tbc	316	50.16	16.559	.932	48.33	51.99	6 85
bronkitis	92	52.45	17.200	1.793	48.88	56.01	13 84
pneumonia	3	63.33	3.512	2.028	54.61	72.06	60 67
Total	483	52.51	17.097	.778	50.98	54.03	6 93

Tabel 17 Hasil uji homogenitas variabel umur

Levene Statistic	df1	df2	Sig.
2.188	3	479	.089

Nilai signifikansi Levene Statistic pada variabel umur > 0.05 , yaitu 0.089, hal ini dapat dianggap bahwa variabel umur pada setiap jenis penyakit berasal dari populasi yang homogen.

Variabel Berat Badan

Nilai varians (standar deviasi) variabel berat badan pada setiap jenis penyakit ditunjukkan pada tabel 18. Semakin tinggi nilai standar deviasi pada masing jenis penyakit, menunjukkan semakin tinggi pula perbedaan rata-rata pada masing-masing jenis penyakit.

Tabel 18 Deskripsi statistik data variabel berat badan

				95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum Maximum
ppok	72	34.27	5.582	.658	32.96	35.58	25 51
tbc	316	45.67	9.424	.530	44.62	46.71	25 80
bronkitis	26	44.84	12.379	2.428	39.84	49.84	25 68
Total	414	43.63	10.045	.494	42.66	44.60	25 80

Tabel 19 Hasil uji homogenitas variabel berat badan

Levene Statistic	df1	df2	Sig.
13.337	2	411	.000

Nilai signifikansi Levene Statistic pada variabel berat badan <0.05 , yaitu 0.00, hal ini dapat dianggap bahwa variabel berat badan pada setiap jenis penyakit tidak berasal dari populasi yang homogen.

Variabel Sistole

Nilai varians (standar deviasi) variabel sistole pada setiap jenis penyakit ditunjukkan pada tabel 20. Semakin tinggi nilai standar deviasi pada masing jenis penyakit, menunjukkan semakin tinggi pula perbedaan rata-rata pada masing-masing jenis penyakit.

Tabel 20 Deskripsi statistik data variabel sistole

					95% Confidence Interval for Mean		Minimum	Maximum
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound		
ppok	72	126.75	21.442	2.527	121.71	131.79	90	190
tbc	316	120.06	20.821	1.171	117.76	122.37	77	180
bronkitis	92	125.61	23.616	2.462	120.72	130.50	76	250
pneumonia	3	116.00	13.077	7.550	83.52	148.48	107	131
Total	483	122.09	21.577	.982	120.16	124.02	76	250

Tabel 21 Hasil uji homogenitas variabel sistole

Levene Statistic	df1	df2	Sig.
.495	3	479	.686

Nilai signifikansi Levene Statistic pada variabel sistole >0.05 , yaitu 0.686, hal ini dapat dianggap bahwa variabel sistole pada setiap jenis penyakit berasal dari populasi yang homogen.

Variabel Diastole

Nilai varians (standar deviasi) variabel diastole pada setiap jenis penyakit ditunjukkan pada tabel 22. Semakin tinggi nilai standar deviasi pada masing jenis penyakit, menunjukkan semakin tinggi pula perbedaan rata-rata pada masing-masing jenis penyakit.

Tabel 22 Deskripsi statistik data variabel diastole

					95% Confidence Interval for Mean		Minimum	Maximum
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound		
ppok	72	76.72	19.333	2.278	72.18	81.27	44	187
tbc	316	74.38	13.253	.746	72.91	75.85	45	130
bronkitis	92	75.09	13.702	1.429	72.25	77.93	46	133

pneumonia	3	69.33	8.021	4.631	49.41	89.26	61	77
Total	483	74.83	14.368	.654	73.55	76.12	44	187

Tabel 23 Hasil uji homogenitas variabel diastole

Diastole			
Levene Statistic	df1	df2	Sig.
.510	3	479	.675

Nilai signifikansi Levene Statistic pada variabel diastole >0.05 , yaitu 0.675, hal ini dapat dianggap bahwa variabel diastole pada setiap jenis penyakit berasal dari populasi yang homogen.

Berdasarkan hasil uji homogenitas, dapat disimpulkan bahwa variabel umur, sistole, diastole memiliki ragam (variance) yang sama atau berasal dari populasi yang homogen. Sedangkan variabel berat badan dapat disimpulkan tidak memiliki ragam (variance) yang sama atau berasal dari populasi yang tidak homogen.

Tabel 24 kesimpulan hasil uji homogenitas

Nama Variabel	Homogenitas Varian
Umur	✓
Berat Badan	x
Sistole	✓
Diastole	✓

4.1.3 Analysis of Variances (Uji F)

Analisis ragam (uji F) pada penelitian ini dilakukan untuk menguji kesamaan rata-rata antar kelompok variabel diagnosa. Dengan menggunakan metode Anova, pengujian kesamaan rata-rata beberapa kelompok dapat dilakukan lebih cepat dan beresiko mengandung kesalahan yang lebih kecil. Perbedaan uji anova dengan uji-t adalah pada jumlah kelompok yang diuji. Uji-t hanya dapat digunakan untuk menguji maksimal rata-rata 2 (dua) kelompok, sedangkan uji anova dapat menguji rata-rata beberapa kelompok. Uji anova hanya dapat digunakan untuk jenis data parametrik, berdistribusi normal dan bersifat homogen. Data yang dapat dilakukan uji anova adalah data

yang bersifat parametrik. Pada studi kasus penelitian ini, variabel diagnosa yang bersifat parametrik adalah variabel umur, berat badan, sistole, diastole. Sehingga untuk variabel yang bersifat non-parametrik dilakukan uji Kruskal Wallis untuk mengetahui ada tidaknya perbedaan ragam antar kelas jenis penyakit. [18]

Variabel Jenis Kelamin

Tabel 25 memuat hasil uji Kruskal Wallis variabel jenis kelamin pada setiap kelas jenis penyakit paru.

Tabel 25 Hasil uji Kruskal Wallis variabel jenis kelamin

Jenis Kelamin	
Chi-Square	9.160
df	3
Asymp. Sig.	.027

a. Kruskal Wallis Test

b. Grouping Variable: Diagnosa

Data pada variabel jenis kelamin merupakan data nominal sehingga untuk mengetahui adanya perbedaan ragam (rata-rata) data bervariasi jenis kelamin pada masing-masing jenis penyakit menggunakan uji Kruskal Wallis. Hasil uji Kruskal Wallis hanya menghasilkan kesimpulan akhir apakah data bervariasi jenis kelamin memiliki perbedaan ragam atau tidak, yang ditunjukkan pada nilai Asymp. Significant <0.05 . Uji Kruskal Wallis tidak dapat menampilkan nilai perbedaan rata-rata pada masing-masing jenis penyakit. Nilai Asymp. Sig. pada variabel jenis kelamin bernilai <0.05 sehingga dapat disimpulkan bahwa data set bervariasi jenis kelamin antar jenis penyakit jenis paru memiliki perbedaan rata-rata.

Variabel Umur

Tabel 26 memuat hasil uji anova variabel umur pada setiap kelas jenis penyakit paru.

Tabel 26 Hasil uji anova variabel umur

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	9161.073	3	3053.691	11.104	.000
Within Groups	131726.164	479	275.002		
Total	140887.237	482			

Nilai significant (hasil uji anova) pada data set umur sebesar 0.00 (<0.05), hal ini menunjukkan bahwa terdapat perbedaan rata-rata nilai umur antar jenis penyakit paru. Pada tabel perbandingan berikut menampilkan selisih perbedaan rata-rata umur antar jenis penyakit paru. Perbedaan rata-rata umur yang signifikan ditandai dengan symbol (*).

Tabel 27 Perbandingan deskripsi umur pada setiap jenis penyakit

(I) Diagnosa	(J) Diagnosa	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
ppok	tbc	12.255*	2.166	.000	8.00	16.51
	bronkitis	9.971*	2.609	.000	4.84	15.10
	pneumonia	-.917	9.772	.925	-20.12	18.28
tbc	ppok	-12.255*	2.166	.000	-16.51	-8.00
	bronkitis	-2.284	1.965	.246	-6.14	1.58
	pneumonia	-13.172	9.620	.172	-32.07	5.73
bronkitis	ppok	-9.971*	2.609	.000	-15.10	-4.84
	tbc	2.284	1.965	.246	-1.58	6.14
	pneumonia	-10.888	9.729	.264	-30.00	8.23
pneumonia	ppok	.917	9.772	.925	-18.28	20.12
	tbc	13.172	9.620	.172	-5.73	32.07
	bronkitis	10.888	9.729	.264	-8.23	30.00

*. The mean difference is significant at the 0.05 level.

Variabel Berat Badan

Tabel 28 memuat hasil uji anova variabel berat badan pada setiap kelas jenis penyakit paru.

Tabel 28 Hasil uji anova variabel berat badan

	Sum of Squares	df	Mean Square	F	Sig.
--	----------------	----	-------------	---	------

Between Groups	7654.574	2	3827.287	46.238	.000
Within Groups	34019.617	411	82.773		
Total	41674.191	413			

Nilai significant (hasil uji anova) pada data set berat badan sebesar 0.00 (<0.05), hal ini menunjukkan bahwa terdapat perbedaan rata-rata nilai berat badan antar jenis penyakit paru. Pada tabel perbandingan berikut menampilkan selisih perbedaan rata-rata berat badan antar jenis penyakit paru. Perbedaan rata-rata berat badan yang signifikan ditandai dengan symbol (*).

Tabel 29 Perbandingan deskripsi berat badan pada setiap jenis penyakit

		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
(I) Diagnosa	(J) Diagnosa				Lower Bound	Upper Bound
ppok	tbc	-11.395*	1.188	.000	-13.73	-9.06
	bronkitis	-10.571*	2.082	.000	-14.66	-6.48
tbc	ppok	11.395*	1.188	.000	9.06	13.73
	bronkitis	.824	1.856	.657	-2.82	4.47
bronkitis	ppok	10.571*	2.082	.000	6.48	14.66
	tbc	-.824	1.856	.657	-4.47	2.82

*. The mean difference is significant at the 0.05 level.

Variabel Sistol

Tabel 30 memuat hasil uji anova variabel sistole pada setiap kelas jenis penyakit paru.

Tabel 30 Hasil uji anova variabel sistole

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4117.393	3	1372.464	2.984	.031
Within Groups	220289.978	479	459.896		
Total	224407.372	482			

Nilai significant (hasil uji anova) pada data set sistole sebesar 0.031 (<0.05), hal ini menunjukkan bahwa terdapat perbedaan rata-rata nilai sistole antar jenis penyakit paru. Pada tabel perbandingan berikut menampilkan selisih perbedaan rata-rata

sistole antar jenis penyakit paru. Perbedaan rata-rata sistole yang signifikan ditandai dengan symbol (*).

Tabel 31 Perbandingan deskripsi sistole pada setiap jenis penyakit

(I) Diagnosa	(J) Diagnosa	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
ppok	tbc	6.688*	2.801	.017	1.19	12.19
	bronkitis	1.136	3.374	.737	-5.49	7.77
	pneumonia	10.750	12.637	.395	-14.08	35.58
tbc	ppok	-6.688*	2.801	.017	-12.19	-1.19
	bronkitis	-5.552*	2.541	.029	-10.54	-.56
	pneumonia	4.062	12.440	.744	-20.38	28.51
bronkitis	ppok	-1.136	3.374	.737	-7.77	5.49
	tbc	5.552*	2.541	.029	.56	10.54
	pneumonia	9.614	12.582	.445	-15.11	34.34
pneumonia	ppok	-10.750	12.637	.395	-35.58	14.08
	tbc	-4.062	12.440	.744	-28.51	20.38
	bronkitis	-9.614	12.582	.445	-34.34	15.11

*. The mean difference is significant at the 0.05 level.

Variabel Diastole

Tabel 32 memuat hasil uji anova variabel diastole pada setiap kelas jenis penyakit paru.

Tabel 32 Hasil uji anova variabel diastole

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	418.374	3	139.458	.674	.568
Within Groups	99080.936	479	206.850		
Total	99499.309	482			

Nilai significant (hasil uji anova) pada data set umur sebesar 0.568 (>0.05), hal ini menunjukkan bahwa tidak terdapat perbedaan rata-rata yang signifikan pada variabel diastole antar jenis penyakit paru. Pada tabel perbandingan berikut menampilkan selisih perbedaan rata-rata umur antar jenis penyakit paru.

Tabel 33 Perbandingan deskripsi diastole pada setiap jenis penyakit

(I) Diagnosa	(J) Diagnosa	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
ppok	tbc	2.341	1.878	.213	-1.35	6.03
	bronkitis	1.630	2.263	.472	-2.82	6.08
	pneumonia	7.389	8.475	.384	-9.26	24.04
tbc	ppok	-2.341	1.878	.213	-6.03	1.35
	bronkitis	-.711	1.704	.677	-4.06	2.64
	pneumonia	5.048	8.343	.545	-11.35	21.44
bronkitis	ppok	-1.630	2.263	.472	-6.08	2.82
	tbc	.711	1.704	.677	-2.64	4.06
	pneumonia	5.759	8.438	.495	-10.82	22.34
pneumonia	ppok	-7.389	8.475	.384	-24.04	9.26
	tbc	-5.048	8.343	.545	-21.44	11.35
	bronkitis	-5.759	8.438	.495	-22.34	10.82

Variabel Batuk

Tabel 34 memuat hasil uji Kruskal Wallis variabel batuk pada setiap kelas jenis penyakit paru.

Tabel 34 Hasil uji Kruskal Wallis variabel batuk

Chi-Square	20.835
df	3
Asymp. Sig.	.000

a. Kruskal Wallis Test

b. Grouping Variable: Diagnosa

Data pada variabel batuk merupakan data ordinal (kategorik) sehingga untuk mengetahui adanya perbedaan ragam (rata-rata) data bervariasi batuk pada masing-masing jenis penyakit menggunakan uji Kruskal Wallis. Hasil uji Kruskal Wallis hanya menghasilkan kesimpulan akhir apakah data bervariasi batuk memiliki perbedaan ragam atau tidak, yang ditunjukkan pada nilai Asymp. Significant <0.05 . Uji Kruskal Wallis tidak dapat menampilkan nilai perbedaan rata-rata pada masing-masing jenis penyakit. Nilai Asymp. Sig. pada variabel batuk bernilai <0.05 sehingga dapat disimpulkan bahwa data set

bervariabel batuk antar jenis penyakit jenis paru memiliki perbedaan rata-rata.

Variabel Sesak

Tabel 35 memuat hasil uji Kruskal Wallis variabel sesak pada setiap kelas jenis penyakit paru.

Tabel 35 Hasil uji Kruskal Wallis variabel sesak

Chi-Square	31.598
df	3
Asymp. Sig.	.000

a. Kruskal Wallis Test

b. Grouping Variable: Diagnosa

Data pada variabel sesak merupakan data ordinal (kategorik) sehingga untuk mengetahui adanya perbedaan ragam (rata-rata) data bervariabel sesak pada masing-masing jenis penyakit menggunakan uji Kruskal Wallis. Hasil uji Kruskal Wallis hanya menghasilkan kesimpulan akhir apakah data bervariabel sesak memiliki perbedaan ragam atau tidak, yang ditunjukkan pada nilai Asymp. Significant <0.05 . Uji Kruskal Wallis tidak dapat menampilkan nilai perbedaan rata-rata pada masing-masing jenis penyakit. Nilai Asymp. Sig. pada variabel sesak bernilai <0.05 sehingga dapat disimpulkan bahwa data set bervariabel sesak antar jenis penyakit jenis paru memiliki perbedaan rata-rata.

Berdasarkan hasil uji Anova dan Kruskal Wallis, dapat disimpulkan bahwa variabel jenis kelamin, umur, sistole, batuk dan sesak memiliki perbedaan rata-rata dan ragam antar kelas penyakit yang cukup signifikan. Sedangkan variabel diastole dapat disimpulkan tidak memiliki perbedaan rata-rata dan ragam antar kelas penyakit yang cukup signifikan.

Tabel 36 Kesimpulan hasil uji anova dan kruskal wallis

Nama Variabel	Perbedaan Ragam	
	Anova	Kruskal Wallis

Jenis Kelamin		√
Umur	√	
Berat Badan	√	
Sistole	√	
Diastole	x	
Batuk		√
Sesak		√

4.1.4 Uji Korelasi

Pada tahap ini, masing-masing variabel yang digunakan untuk mendiagnosa jenis penyakit paru-paru dihitung nilai korelasinya terhadap variabel target (jenis penyakit paru) untuk mengetahui keeratan hubungan antara masing-masing variabel dengan variabel target, serta untuk mengetahui arah hubungan yang terjadi.

Tabel 37 Hasil Uji Korelasi

Nama Variabel	Nilai Korelasi	Keterangan
Jenis Kelamin	0.138	Sangat Lemah
Umur	-0.132	Sangat Lemah
Berat Badan	0.359	Lemah
Sistole	-0.008	Sangat Lemah
Diastole	-0.035	Sangat Lemah
Batuk	0.19	Sangat Lemah
Sesak	-0.186	Sangat Lemah

Dari hasil pengujian nilai korelasi pada masing-masing variabel diagnosa terhadap variabel target (jenis penyakit paru), dapat disimpulkan bahwa seluruh variabel diagnosa yang digunakan pada penelitian ini memiliki keeratan hubungan yang rendah terhadap variabel target atau memiliki nilai korelasi <0.4 .

Nilai korelasi yang semakin mendekati nilai 1 atau -1 menunjukkan bahwa hubungan antara variabel diagnosa dengan variabel jenis penyakit paru (target) sangat kuat, sedangkan apabila nilai korelasi mendekati nol (0) berarti hubungan antara variabel diagnosa tersebut dengan variabel target sangat lemah. Korelasi bernilai positif menunjukkan hubungan kedua variabel tersebut searah (apabila nilai variabel diagnosa tinggi maka nilai kategori jenis penyakit juga tinggi) sedangkan apabila

korelasi bernilai negatif menunjukkan bahwa hubungan kedua variabel tersebut terbalik (apabila nilai variabel diagnosa tinggi maka nilai kategori jenis penyakit rendah).

4.1.5 Karakteristik Data

Tabel 38 menunjukkan karakteristik data set pada masing-masing variabel diagnosa berdasarkan hasil uji hipotesa statistika yang telah dilakukan pada tahapan sebelumnya, yaitu uji normalitas, uji homogenitas, *Analysis of Variance* (uji F) dan uji korelasi.

Tabel 38 Statistika deskriptif variabel titik embun

Nama Variabel	Jenis Data	Normalitas	Homogenitas Varian	Perbedaan Ragam	
				Anova	Kruskal Wallis
Jenis Kelamin	Nominal	√			√
Umur	Rasio	√	√	√	
Berat Badan	Rasio	√	x	√	
Sistole	Rasio	√	√	√	
Diastole	Rasio	√	√	x	
Batuk	Ordinal	√			√
Sesak	Ordinal	√			√

4.1.5.1 Variabel Jenis Kelamin

Data set bervariasi jenis kelamin memiliki jenis data nominal karena terdiri dari kategori-kategori yang saling lepas (tidak dapat terjadi bersamaan) dan tidak memiliki hubungan prioritas atau peringkat (lebih dari) antar kategori tersebut, yaitu jenis kelamin laki-laki dan perempuan. Pada penelitian ini, kategori pada variabel jenis kelamin yang bersifat karakter (huruf) diubah menjadi kategori numerik (angka) untuk memudahkan pemrosesan data pada uji statistik maupun dalam proses pembuatan model ANN. Jenis kelamin perempuan diubah menjadi nilai 1 (satu) sedangkan laki-laki diubah menjadi nilai 0 (nol).

Berdasarkan uji normalitas, persebaran data pada variabel jenis kelamin berdistribusi normal. Data set bervariasi jenis kelamin

tidak dapat dilakukan uji homogenitas karena jenis data variabel jenis kelamin adalah non-parametrik. Sehingga untuk mengetahui apakah terdapat perbedaan rata-rata (ragam) jenis kelamin antar jenis penyakit paru, peneliti menggunakan metode Kruskal Wallis. Dari hasil uji Kruskal Wallis dapat disimpulkan bahwa terdapat perbedaan rata-rata (ragam) jenis kelamin antar jenis penyakit paru (PPOK, tuberculosis, bronkitis, pneumonia).

4.1.5.2 Variabel Umur

Data set bervariasi umur memiliki jenis data rasio karena terdiri dari nilai-nilai yang saling lepas (tidak dapat terjadi bersamaan) dan memiliki nilai 0 (nol) yang berarti. Berdasarkan uji normalitas, dapat disimpulkan bahwa persebaran data pada variabel umur berdistribusi normal. Berdasarkan uji homogenitas, data set bervariasi umur juga memiliki kesamaan varian (bersifat homogen). Dan berdasarkan uji anova (uji F), dapat disimpulkan bahwa terdapat perbedaan rata-rata (ragam) umur yang signifikan pada beberapa jenis penyakit paru (PPOK, tuberculosis, bronkitis).

4.1.5.3 Variabel Berat Badan

Data set bervariasi berat badan memiliki jenis data rasio karena terdiri dari nilai-nilai yang saling lepas (tidak dapat terjadi bersamaan) dan memiliki nilai 0 (nol) yang berarti. Berdasarkan uji normalitas, dapat disimpulkan bahwa persebaran data pada variabel berat badan berdistribusi normal. Berdasarkan uji homogenitas, data set bervariasi berat badan tidak memiliki kesamaan varian (bersifat homogen). Tetapi berdasarkan uji anova (uji F), terdapat perbedaan rata-rata (ragam) berat badan yang signifikan pada beberapa jenis penyakit paru (PPOK, tuberculosis, bronkitis).

4.1.5.4 Variabel Sistole

Data set bivariabel sistole memiliki jenis data rasio karena terdiri dari nilai-nilai yang saling lepas (tidak dapat terjadi bersamaan) dan memiliki nilai 0 (nol) yang berarti. Berdasarkan uji normalitas, dapat disimpulkan bahwa persebaran data pada variabel sistole berdistribusi normal. Berdasarkan uji homogenitas, data set bivariabel sistole juga memiliki kesamaan varian (bersifat homogen). Dan berdasarkan uji anova (uji F), dapat disimpulkan bahwa terdapat perbedaan rata-rata (ragam) sistole yang signifikan pada beberapa jenis penyakit paru (PPOK, tuberculosis, bronkitis).

4.1.5.5 Variabel Diastole

Data set bivariabel diastole memiliki jenis data rasio karena terdiri dari nilai-nilai yang saling lepas (tidak dapat terjadi bersamaan) dan memiliki nilai 0 (nol) yang berarti. Berdasarkan uji normalitas, dapat disimpulkan bahwa persebaran data pada variabel diastole berdistribusi normal. Berdasarkan uji homogenitas, data set bivariabel diastole juga memiliki kesamaan varian (bersifat homogen). Tetapi berdasarkan uji anova (uji F), dapat disimpulkan bahwa tidak terdapat perbedaan rata-rata (ragam) diastole antar jenis penyakit paru (PPOK, tuberculosis, bronkitis, pneumonia).

4.1.5.6 Variabel Batuk

Data set bivariabel batuk memiliki jenis data ordinal karena terdiri dari kategori-kategori yang saling lepas (tidak dapat terjadi bersamaan) dan memiliki hubungan prioritas atau peringkat (lebih dari) antar kategori tersebut, yaitu 'batuk', 'tidak batuk', dan 'batuk berdarah'. Pada penelitian ini, kategori pada variabel batuk yang bersifat karakter (huruf) diubah menjadi kategori numerik (angka) untuk memudahkan pemrosesan data pada uji statistik maupun dalam proses pembuatan model ANN. Status penderita yang tidak mengalami gejala batuk nilainya diubah menjadi nilai 0 (nol), penderita

yang mengalami gejala batuk nilainya diubah menjadi nilai 1 (satu), dan penderita yang mengalami gejala batuk disertai darah nilainya diubah menjadi 2 (dua).

Berdasarkan uji normalitas, persebaran data pada variabel batuk berdistribusi normal. Data set bervariasi variabel batuk tidak dapat dilakukan uji homogenitas karena jenis data variabel batuk adalah non-parametrik. Sehingga untuk mengetahui apakah terdapat perbedaan rata-rata (ragam) status batuk antar jenis penyakit paru, peneliti menggunakan metode Kruskal Wallis. Dari hasil uji Kruskal Wallis dapat disimpulkan bahwa terdapat perbedaan rata-rata (ragam) status batuk antar jenis penyakit paru (PPOK, tuberculosis, bronkitis, pneumonia).

4.1.5.7 Variabel Sesak

Data set bervariasi variabel sesak memiliki jenis data ordinal karena terdiri dari kategori-kategori yang saling lepas (tidak dapat terjadi bersamaan) dan memiliki hubungan prioritas atau peringkat (lebih dari) antar kategori tersebut, yaitu 'sesak' dan 'tidak sesak'. Pada penelitian ini, kategori pada variabel sesak yang bersifat karakter (huruf) diubah menjadi kategori numerik (angka) untuk memudahkan pemrosesan data pada uji statistik maupun dalam proses pembuatan model ANN. Status penderita yang tidak mengalami gejala sesak nilainya diubah menjadi nilai 0 (nol) sedangkan penderita yang mengalami gejala sesak nilainya diubah menjadi nilai 1 (satu).

Berdasarkan uji normalitas, persebaran data pada variabel sesak berdistribusi normal. Data set bervariasi variabel sesak tidak dapat dilakukan uji homogenitas karena jenis data variabel sesak adalah non-parametrik. Sehingga untuk mengetahui apakah terdapat perbedaan rata-rata (ragam) status sesak antar jenis penyakit paru, peneliti menggunakan metode Kruskal Wallis. Dari hasil uji Kruskal Wallis dapat disimpulkan bahwa terdapat perbedaan rata-rata (ragam) status sesak antar jenis penyakit paru (PPOK, tuberculosis, bronkitis, pneumonia).

4.2 Penentuan Model Neural Network

Jenis model ANN yang digunakan pada penelitian ini adalah *feed-forward backpropagation*, dimana sinyal informasi bergerak dari kedua arah (dari layer input ke layer output dan sebaliknya) secara iteratif hingga mencapai titik keseimbangan (ekuilibrium) serta nilai bobot dan bias yang sesuai untuk meminimalkan nilai error (*Mean Squared Error*). Metode *feed-forward backpropagation* memiliki kemampuan untuk mengatasi permasalahan pelatihan klasifikasi dengan skala data yang luas (powerfull). Namun metode *feed-forward backpropagation* memiliki kelemahan dalam kecepatan konvergen yang lambat dan tidak stabil, hal ini dipengaruhi oleh bobot awal yang dipilih secara acak (*random*). Langkah pembelajaran metode *neural network feed-forward backpropagation* adalah sebagai berikut [14]:

Langkah 0:

- Inisiasi nilai bobot dan bias secara acak (dengan nilai bias pada interval -1 dan 1).

Feedforward

Langkah 1:

- Masing-masing node pada input layer menerima sinyal masukan kemudian disebarkan ke node-node hidden layer.

Langkah 2:

- Masing-masing node pada hidden layer dikalikan dengan nilai bobot dan dijumlahkan dengan nilai bias.
- Kemudian menghitung sesuai dengan fungsi aktivasi yang digunakan.

Langkah 3:

- Masing-masing node pada output layer dikalikan dengan nilai bobot dan dijumlahkan dengan nilai bias.
- Kemudian menghitung sesuai dengan fungsi aktivasi yang digunakan.

Backpropagation

Langkah 4:

- Masing-masing node pada output layer menerima pola target sesuai dengan pola masukan (saat pelatihan/training).
- Menghitung nilai error.
- Menghitung perbaikan nilai bobot.
- Menghitung perbaikan nilai bias.

Langkah 5:

- Masing-masing bobot dan bias yang menghubungkan node-node output layer dengan node-node pada hidden layer dikalikan bias kemudian dijumlahkan sebagai masukan ke node-node lapisan berikutnya.
- Menghitung nilai error dengan cara dikalikan dengan turunan dari fungsi aktivasinya.
- Menghitung perbaikan nilai bobot.
- Menghitung perbaikan nilai bias.

Langkah 6:

- Masing-masing node pada output layer diperbaiki nilai bobot dan biasnya.
- Masing-masing node pada hidden layer diperbaiki nilai bobot dan biasnya.

Langkah 7:

- Uji kondisi pemberhentian (akhir iterasi).

4.3 Penentuan Struktur Model ANN

Tabel 39 merupakan rancangan jumlah layer dan node pada masing-masing layer.

Tabel 39 Rancangan Struktur Model ANN

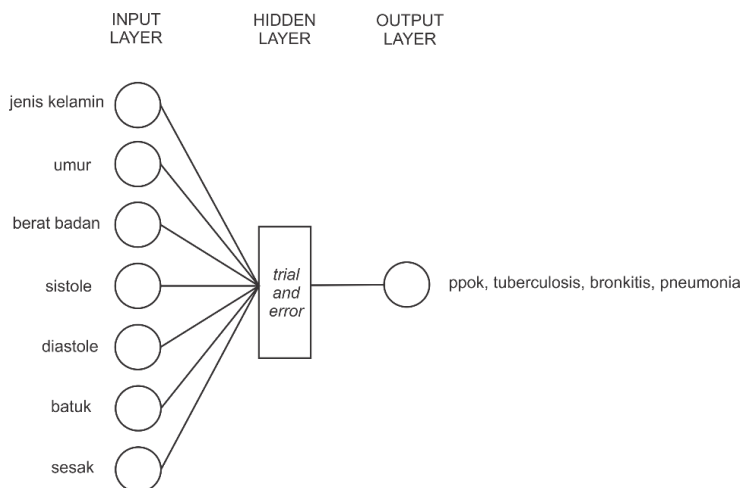
Layer	Jumlah Layer	Jumlah Node	Data
Input	1	5-7	Jenis kelamin, umur, berat badan, sistole, diastole, batuk, sesak.
Hidden	Trial and error		Jumlah node, fungsi aktivasi
Output	1	1 atau 4	PPOK, tuberculosis, bronkitis, pneumonia.

4.3.1 Penentuan Input Layer

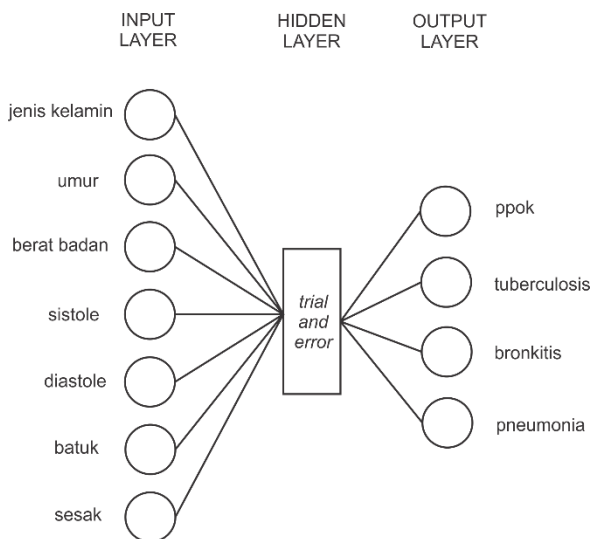
Data yang digunakan sebagai masukan (input) adalah data yang yang diperlukan untuk mendiagnosa jenis penyakit paru. Data input yang digunakan pada penelitian ini sebanyak 7 (tujuh) variabel, yaitu data bervariasi jenis kelamin, umur, berat badan, sistole, diastole, batuk, dan sesak. Dalam penelitian ini juga akan membandingkan validitas model dengan 5 (lima) hingga 7 (tujuh) variabel yang memiliki nilai korelasi (R) tertinggi berdasarkan hasil uji korelasi. Sehingga pada penelitian ini jumlah node pada input layer adalah sebanyak 5 (lima) hingga 7 (tujuh) variabel.

4.3.2 Penentuan Output Layer

Output yang diharapkan pada penelitian ini adalah model ANN dapat mendiagnosa 4 jenis penyakit paru (ppok, tuberculosis, bronkitis, pneumonia). Pada penelitian ini akan membandingkan validitas model dengan 4 (empat) dan 1 (satu) buah node pada output layer. Masing-masing node pada model dengan 4 (empat) buah node akan memuat 1 jenis penyakit, sedangkan model dengan 1 (satu) buah node akan memuat 4 jenis penyakit.



Gambar 11 Struktur ANN dengan 1 variabel input



Gambar 12 Struktur ANN dengan 4 variabel input

4.3.3 Penentuan Hidden Layer

Tabel 40 berikut menampilkan nilai Mean Squared Error (MSE) dan korelasi (R) hasil uji data training terhadap variasi jumlah node pada setiap layer dan jumlah layer pada hidden layer. Fungsi aktivasi yang digunakan pada hidden layer dan output layer masing-masing adalah tansig dan purelin. Dan fungsi pembelajaran yang digunakan adalah trainlm. Dari hasil training, dapat disimpulkan bahwa jumlah node pada hidden layer mempengaruhi nilai MSE dan korelasi (R), semakin tinggi jumlah node pada hidden layer maka semakin rendah nilai MSE dan semakin tinggi nilai korelasinya. Sedangkan jumlah hidden layer yang digunakan tidak mempengaruhi nilai MSE dan nilai korelasi (R).

Tabel 40 Struktur Model ANN dengan Fungsi Pembelajaran Trainlm

Jumlah Hidden Layer	Jumlah Node	MSE	R (%)			
			Training	Validation	Test	All
1	1	0.121	73.2	60.6	77.6	71.9
	2	0.102	74.3	67.6	65.6	72
	3	0.089	75.6	72.6	75.6	75.2
	4	0.079	83	76.4	70.4	80.2
	5	0.060	80.7	81.4	77.8	80.4
	6	0.132	81.1	62.4	77.9	77.3
	7	0.077	80.9	76.5	77.9	79.8
	8	0.085	80.5	74.3	83.7	80
	9	0.070	81.8	79	74.7	80.3
	10	0.086	86.8	74.7	59.9	79.7
2	[1 1]	0.061	71	82.3	73.1	73.1
	[2 2]	0.060	78.5	82.3	81.7	79.5
	[3 3]	0.072	81.7	78.3	77.6	80.6
	[4 4]	0.057	79.1	83.3	82.7	80.2
	[5 5]	0.090	81.5	72.5	78.3	79.7
	[6 6]	0.078	81	76.2	60.3	76.4
	[7 7]	0.070	79.3	78.9	77.3	78.9
	[8 8]	0.077	81.9	76.6	76.3	80.2
	[9 9]	0.098	80.8	69.4	57.8	75.8

Jumlah Hidden Layer	Jumlah Node	MSE	R (%)			
			Training	Validation	Test	All
	[10 10]	0.094	82.4	71.2	80.9	80.6

Tabel 41 berikut menampilkan nilai Mean Squared Error (MSE) dan korelasi (R) hasil uji data training terhadap variasi jumlah node pada setiap layer dan jumlah layer pada hidden layer. Fungsi aktivasi yang digunakan pada hidden layer dan output layer masing-masing adalah tansig dan purelin. Dan fungsi pembelajaran yang digunakan adalah trainbr. Dari hasil training, dapat disimpulkan bahwa jumlah node pada hidden layer mempengaruhi nilai MSE dan korelasi (R), semakin tinggi jumlah node pada hidden layer maka semakin rendah nilai MSE dan semakin tinggi nilai korelasinya. Sedangkan jumlah hidden layer yang digunakan tidak mempengaruhi nilai MSE dan nilai korelasi (R).

Tabel 41 Struktur Model ANN dengan Fungsi Pembelajaran Trainbr

Jumlah Hidden Layer	Jumlah Node	MSE	R (%)		
			Training	Test	All
1	1	0.084	74.2	66.9	73.1
	2	0.068	79.7	83.5	80.3
	3	0.058	82.9	67.4	80.6
	4	0.060	82.4	75.2	81.3
	5	0.061	82	82.3	82
	6	0.057	83.4	75.3	82.1
	7	0.063	81.5	83	81.7
	8	0.051	85.1	58.3	80.8
	9	0.053	84.7	74.7	83.1
	10	0.042	88.1	68.6	85.1
2	[1 1]	0.083	74.6	64.1	73.1
	[2 2]	0.065	80.7	77.3	80.2
	[3 3]	0.055	84.1	51.5	78.5
	[4 4]	0.058	82.9	74.3	81.5
	[5 5]	0.058	83.1	74.7	81.9

Jumlah Hidden Layer	Jumlah Node	MSE	R (%)		
			Training	Test	All
	[6 6]	0.058	83.2	74.9	81.9
	[7 7]	0.051	85.5	68.4	82.9
	[8 8]	0.058	83	80.8	82.7
	[9 9]	0.051	85.1	53.9	79.7
	[10 10]	0.049	85.7	76.3	84.1

4.4 Penentuan Parameter

4.4.1 Fungsi Aktivasi

Fungsi aktivasi merupakan fungsi matematis yang digunakan untuk menghasilkan output yang sesuai dengan target dari input yang diberikan. Berikut ini merupakan fungsi aktivasi yang terdapat pada model *neural network feed-forward backpropagation*.

- Logsig (sigmoid unipolar/biner): digunakan untuk metode *backpropagation*. Fungsi ini memiliki nilai output yang terletak pada interval 0 hingga 1.
- Purelin (identitas/linier): fungsi purelin mampu memprediksi nilai yang diluar rentang data pelatihan. Fungsi ini memiliki nilai output yang sama dengan nilai inputnya.
- Tansig (sigmoid bipolar): fungsi ini hampir sama dengan fungsi logsig, tetapi fungsi ini memiliki nilai output yang terletak pada interval 1 hingga -1.

Tabel 42 berikut menampilkan nilai Mean Squared Error (MSE) dan korelasi (R) hasil uji data training terhadap variasi penggunaan fungsi aktivasi pada hidden layer dan output layer dengan menggunakan 7 node input (jenis kelamin, umur, berat badan, sistole, diastole, batuk, sesak), 1 hidden layer dengan 1 node serta 1 node output (ppok, tuberculosis, bronkitis, pneumonia). Dari hasil training, dapat disimpulkan bahwa penggunaan fungsi aktivasi logsig pada output layer menghasilkan nilai MSE paling tinggi (>50%) dibanding penggunaan fungsi aktivasi purelin atau tansig.

Tabel 42 Variasi Fungsi Aktivasi pada 1 Node Output

Fungsi Aktivasi		Fungsi Training	MSE	R (%)		
Hidden Layer	Output Layer			Training	Testing	All
logsig	logsig	trainbr	0.587	0	0	0
	purelin		0.216	60.9	63	61.7
	tansig		0.228	61.5	64.7	61.9
purelin	logsig		0.507	63	55.4	61.9
	purelin		0.255	51.7	46.9	50.8
	tansig		0.262	52.8	50.9	52.6
tansig	logsig		0.527	63.3	60.8	62.2
	purelin		0.224	62.3	60.4	62
	tansig		0.222	62.3	59.8	61.8

Sedangkan tabel 43 berikut menampilkan nilai *Mean Squared Error* (MSE) dan korelasi (R) hasil uji data training terhadap variasi penggunaan fungsi aktivasi pada hidden layer dan output layer dengan menggunakan 7 node input (jenis kelamin, umur, berat badan, sistole, diastole, batuk, sesak), 1 hidden layer dengan 1 node serta 4 node output (ppok, tuberculosi, bronkitis, pneumonia). Dari hasil *training*, dapat disimpulkan bahwa penggunaan fungsi aktivasi logsig pada output layer juga menghasilkan nilai MSE paling tinggi (>20%) dibanding penggunaan fungsi aktivasi purelin atau tansig.

Tabel 43 Variasi Fungsi Aktivasi pada 1 Node Output

Fungsi Aktivasi		Fungsi Training	MSE	R (%)		
Hidden Layer	Output Layer			Training	Testing	All
logsig	logsig	trainbr	0.245	51.7	65.4	53.7
	purelin		0.086	73.4	71.4	73
	tansig		0.076	76.9	67.4	75.4
purelin	logsig		0.273	54.1	53	53.9
	purelin		0.087	70	75.4	70.9
	tansig		0.084	74.3	77.6	74.8
tansig	logsig		0.221	73.2	70.3	72.8
	purelin		0.089	72.2	78.6	73.2

	tansig		0.085	73.8	79.2	74.7
--	--------	--	-------	------	------	------

4.4.2 Fungsi Pelatihan

Fungsi pelatihan (*training function*) merupakan fungsi (algoritma) yang digunakan dalam proses pelatihan *neural network* agar dapat dengan mudah dan cepat mengenali pola (target) dalam sekali proses training [16]. Berikut ini tabel variasi fungsi pembelajaran pada neural network.

Tabel 44 Fungsi Pembelajaran Neural Network

Fungsi	Algorithm	Deskripsi
Traingd	<i>Basic Gradient Descent</i>	Respon lambat, dapat digunakan dalam mode latihan tambahan.
Traingdm	<i>Gradient Descent with momentum</i>	Secara umum lebih cepat dibandingkan traingd. Dapat digunakan dalam mode latihan tambahan.
Traingdx	<i>Variable Learning Rate Backpropagation</i>	Tingkat pembelajaran yang adaptive. Pembelajaran lebih cepat dibandingkan traingd, tetapi hanya dapat digunakan dalam sekumpulan mode pembelajaran.
Trainrp	<i>Resilient Backpropagation</i>	Algoritma sekumpulan mode pembelajaran yang sederhana dengan pemusatan yang cepat dan kebutuhan penyimpanan yang minimal.
Traincgf	<i>Fletcher-Powell Conjugate Gradient</i>	Membutuhkan penyimpanan yang paling kecil dari algoritma gradien yang dihubungkan.
Traincgp	<i>Polak-Ribiere Conjugate Gradient</i>	Membutuhkan sedikit lebih banyak penyimpanan daripada traincgf. Pemusatan dari beberapa permasalahan jauh lebih cepat.
Traincgb	<i>Conjugate Gradient</i>	Membutuhkan lebih banyak penyimpanan daripada traincgp.

Fungsi	Algorithm	Deskripsi
	<i>with Powell/Beale Restarts</i>	Pada umumnya, pemusatan jauh lebih cepat.
Trainscg	<i>Scaled Conjugate Gradient</i>	Hanya algoritma gradien konjugasi yang tidak membutuhkan garis pencarian. Sangat bagus untuk algoritma pelatihan dengan tujuan yang umum.
Trainbfg	<i>BFGS Quasi-Newton</i>	Membutuhkan penyimpanan dari matriks Hessian dan memiliki lebih banyak komputasi dalam masing-masing iterasi daripada algoritma gradien konjugasi, tetapi biasanya terpusat dalam interaksi yang lebih sedikit.
Trainoss	<i>One-Step Secant</i>	Menggabungkan metode gradien konjugasi dan metode quasi-Newton.
Trainlm	<i>Levenberg Marquardt</i>	Algoritma pelatihan sangat cepat untuk jaringan dengan ukuran medium. Memiliki fitur pengurangan penyimpanan yang digunakan ketika setelah pelatihan memiliki ukuran yang besar.
Trainbr	<i>Bayesian Regularization</i>	Memodifikasi algoritma pelatihan Levenberg-Marquardt untuk menghasilkan jaringan yang lebih baik. Mengurangi kesulitan dalam menentukan arsitektur jaringan yang optimum.

Tabel 45 berikut menampilkan nilai Mean Squared Error (MSE) dan korelasi (R) hasil uji data training terhadap variasi penggunaan fungsi pembelajaran pada neural network dengan menggunakan 7 node input (jenis kelamin, umur, berat badan, sistole, diastole, batuk, sesak), 1 hidden layer dengan 1 node

serta 4 node output (ppok, tuberculosis, bronkitis, pneumonia). Fungsi aktivasi yang digunakan pada hidden layer dan output layer masing-masing adalah tansig dan purelin. Dari hasil training, dapat disimpulkan bahwa penggunaan fungsi pembelajaran **trainlm** dan **trainbr** menghasilkan nilai MSE lebih rendah (<10%) dan nilai korelasi (R) paling tinggi dibanding penggunaan fungsi pembelajaran yang lain.

Tabel 45 Penentuan Fungsi Pembelajaran

Fungsi Pembelajaran	MSE	R (%)			
		Training	Validation	Testing	All
Traingd	0.152	50.9	43.5	47.1	49.3
Traingdm	0.136	43.9	55.4	48.7	46.3
Traingdx	0.088	72.9	72.6	69.6	72.4
Trainrp	0.132	58.3	54.9	47.8	56.2
Traincgf	0.131	59.6	54.9	54.1	58.1
Traincgp	0.087	73.8	73.1	63.2	72.1
Traincgb	0.093	70.9	70.9	81.2	72.5
Trainscg	0.100	76.4	68.2	59.6	72.7
Trainbfg	0.113	72.9	63.3	73.5	71.5
Trainoss	0.092	59.1	72.2	54.9	60.4
Trainlm	0.059	72.3	83	67.1	73.1
Trainbr	0.088	72.7		76.3	73.2

4.4.3 Jumlah Epoch

Epoch adalah suatu jangka waktu (satu set putaran) yang diperlukan untuk pelatihan sebuah ANN sehingga kesalahan mendekati nol (0). Jumlah epoch juga merupakan salah satu faktor yang dapat menentukan proses pemberhentian iterasi pada setiap kali *training* maupun *testing*. Untuk menghasilkan model yang optimal, diperlukan proses *training* atau *testing* dengan jumlah iterasi yang optimal, yaitu tidak terlalu sedikit dan tidak terlalu banyak. Oleh sebab itu, jumlah epoch dapat mempengaruhi performa model yang dihasilkan. Jumlah epoch pada penelitian ini dalam setiap kali training dan testing adalah sebanyak 1000.

4.4.4 Fungsi Performa

Fungsi performa pada neural network digunakan sebagai indikator untuk uji akurasi dari hasil prediksi suatu model. Mean Squared Error (MSE) merupakan indikator untuk mengukur kesalahan dengan cara menghitung rata-rata kuadrat selisih antara nilai aktual dengan nilai yang diprediksi (error) [19]. Fungsi performa MSE pada umumnya digunakan untuk mengukur akurasi dari teknik peramalan dan prediksi (klasifikasi). Semakin rendah nilai MSE maka semakin tinggi akurasi prediksi yang dihasilkan.

$$MSE = \frac{\sum (Aktual - forecast)^2}{n-1}, \text{ dimana } n = \text{banyaknya data}$$

4.4.5 Fungsi Pembelajaran

Fungsi pembelajaran (*adaption learning function*) merupakan fungsi (algoritma) yang digunakan dalam inisiasi (adaptasi) proses pelatihan (*training*). Fungsi pembelajaran digunakan untuk menghitung perbaikan nilai bobot secara teratur dan terus menerus [20]. Pada model *feed-forward backpropagation neural network* ada 2 jenis, yaitu *learnngd* dan *learnngdm*. *Learnngd* menggunakan metode *gradient descent*, cocok digunakan untuk turunan model *backpropagation*. Sedangkan *learnngdm* menggunakan metode *gradient descent momentum*, cocok digunakan untuk variasi momentum model *backpropagation*.

Pada penelitian ini, fungsi pembelajaran yang digunakan adalah *learnngdm*. Berdasarkan hasil percobaan yang telah dilakukan, fungsi pembelajaran *learnngdm* menghasilkan nilai MSE yang lebih rendah dan nilai korelasi yang lebih tinggi dibandingkan fungsi pembelajaran *learnngd*.

Halaman ini sengaja dikosongkan

BAB V

IMPLEMENTASI

5.1 Pemodelan ANN dengan MATLAB

Pembuatan model ANN dilakukan dengan bantuan perangkat *toolbox neural network* (nntool) pada Matlab version 7.14 (R2012a) dan Neural Network Toolbox version 7.0.3 (R2012a). Pada tahap ini akan dilakukan proses training dan testing beberapa model ANN dengan kriteria data (variabel input) dan model (struktur dan parameter) yang telah ditentukan pada tahapan sebelumnya untuk menghasilkan model ANN yang optimal (memiliki nilai MSE <10% dan nilai korelasi >80%) dalam mendiagnosa jenis penyakit paru pada studi kasus penelitian ini.

Berdasarkan tahapan persiapan data, penentuan jenis model, struktur model ANN dan parameternya, pemodelan ANN akan dilakukan dengan kriteria sebagai berikut:

5.1.1 Memasukkan Data

Tahapan pertama dalam pembangunan model ANN adalah memasukkan data training dan data testing ke dalam jendela workspace Matlab. Data training terdiri dari data variabel input dan data targetan keluaran yang diharapkan. Sedangkan data testing hanya terdiri dari data variabel input. Langkah untuk memasukkan data pada workspace pada Matlab dapat dilakukan melalui Graphic User Interface (GUI) maupun script pada jendela command.

5.1.2 Membuat Model ANN

Tahap kedua dalam pembangunan model ANN adalah pembuatan model sesuai dengan struktur dan parameter yang telah dirancang pada tahap perancangan. Langkah untuk membuat model ANN pada Matlab dapat dilakukan melalui Graphic User Interface (GUI) maupun script pada jendela command.

5.1.3 Melakukan Training

Tahap training pada penelitian ini dilakukan untuk mencari model ANN paling optimal dalam mendiagnosa penyakit paru, yaitu model yang memiliki nilai korelasi yang tinggi antara input dengan output, dan nilai MSE yang rendah. Langkah untuk melakukan training model ANN pada Matlab dapat dilakukan melalui Graphic User Interface (GUI) maupun script pada jendela command.

5.1.4 Melakukan Testing

Tahap uji coba (*testing*) pada penelitian ini dilakukan untuk mensimulasikan model paling optimal (hasil proses *training*) dengan data masukan yang berbeda dari proses *training* (data *testing*). Pada proses *testing*, model akan menghasilkan nilai keluaran berupa prediksi (diagnosa) jenis penyakit paru-paru. Nilai keluaran tersebut kemudian diklasifikasikan ke dalam 4 jenis penyakit (ppok, tuberculosis, bronkitis, pneumonia).

Berdasarkan data *sample* yang diambil dari studi kasus penelitian ini, urutan penempatan kelas penyakit paru, untuk 1 node maupun 4 node pada output layer ditentukan berdasarkan urutan tingkat keseriusan jenis penyakit paru, dari tingkat yang paling tinggi hingga tingkat yang paling rendah, yaitu PPOK, tuberculosis, bronkitis, pneumonia. Penentuan diagnosa pada model dengan 1 node pada output layer yang paling sesuai dengan data sample pada studi kasus penelitian ini adalah dengan interval yang ditunjukkan pada tabel 46. Penentuan interval nilai yang digunakan pada model dengan 1 node pada output layer tersebut didapatkan dari hasil analisa hasil prediksi pada beberapa model ANN agar menghasilkan performa klasifikasi yang tinggi.

Sedangkan output pada model dengan 4 node pada output layer menghasilkan nilai dari interval 0-1 (nol hingga satu). Untuk penentuan diagnosa pada model dengan 4 node pada output layer ditentukan berdasarkan kelas (node) yang menghasilkan

persentase nilai prediksi paling tinggi. Kemudian nilai output yang memiliki nilai presentase paling tinggi akan dibulatkan (konversi) ke nilai 1. Sedangkan 3 (tiga) nilai output terendah lainnya dibulatkan menjadi 0. Contoh proses penentuan hasil diagnosa pada model dengan 4 node pada output layer ditunjukkan pada tabel 47.

Tabel 46 Klasifikasi nilai keluaran dengan 1 node pada output layer

Nilai Output	Hasil Diagnosa
<1.54	PPOK
1.55 – 2.54	Tuberculosis
2.55 – 3.54	Bronkitis
>3.55	Pneumonia

Tabel 47 Contoh penentuan diagnosa dengan 4 node pada output layer

Nilai Output	Konversi	Hasil Diagnosa
0.0795	0	Tuberculosis
0.7795	1	
0.1428	0	
0.0018	0	

5.2 Validasi Model

Uji validitas model ANN dilakukan untuk menguji keabsahan model ANN yang dihasilkan dari proses training [15]. Apabila model tersebut tetap memprediksi dengan akurasi yang tinggi ketika dilakukan uji coba (testing) dengan data masukan yang berbeda dari data masukan pada proses training, maka model tersebut dapat dikatakan valid (sah). Model yang valid akan direkomendasikan sebagai dasar untuk pembuatan sistem diagnosa penyakit paru-paru.

Uji validitas model dilakukan dengan cara menghitung nilai akurasi, precision, dan recall dari hasil prediksi pada proses testing dengan bantuan tabel *contingency (confusion matrix)*.

Halaman ini sengaja dikosongkan

BAB VI

HASIL DAN PEMBAHASAN

6.1 Lingkungan Uji Coba

Lingkungan uji coba merupakan kriteria perangkat yang digunakan dalam melakukan percobaan untuk mendapatkan model ANN yang optimal pada penelitian ini. Lingkungan uji coba pada penelitian ini terdiri dari perangkat keras (hardware) dan perangkat lunak (software). Lingkungan perangkat keras yang digunakan pada penelitian ini ditunjukkan pada tabel 48.

Tabel 48 Lingkungan perangkat keras yang digunakan

Perangkat Keras	Spesifikasi
Jenis	Personal Computer
Processor	Intel Core i7
RAM	4 GB
Hard Disk Drive	1 TB

Sedangkan lingkungan perangkat lunak yang digunakan pada penelitian ini ditunjukkan pada tabel 49.

Tabel 49 Lingkungan perangkat lunak yang digunakan

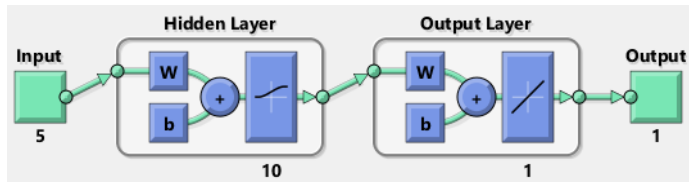
Perangkat Lunak	Fungsi
Windows 8	Sistem Operasi
Matlab 2016	Membentuk model, melakukan training dan testing model
Microsoft Excel 2013	Mengelola data

6.2 Implementasi Model 1 Node Output

Berikut ini model neural network paling optimal berdasarkan hasil training dan testing yang menggunakan 1 (satu) node pada output layer. Model tersebut dibedakan berdasarkan jumlah variabel (node) pada input layer. Pada penelitian ini dilakukan percobaan pembuatan model dengan jumlah variabel input sebanyak 5 (lima) hingga 7 (tujuh) dari jumlah variabel input yang memiliki nilai korelasi tertinggi.

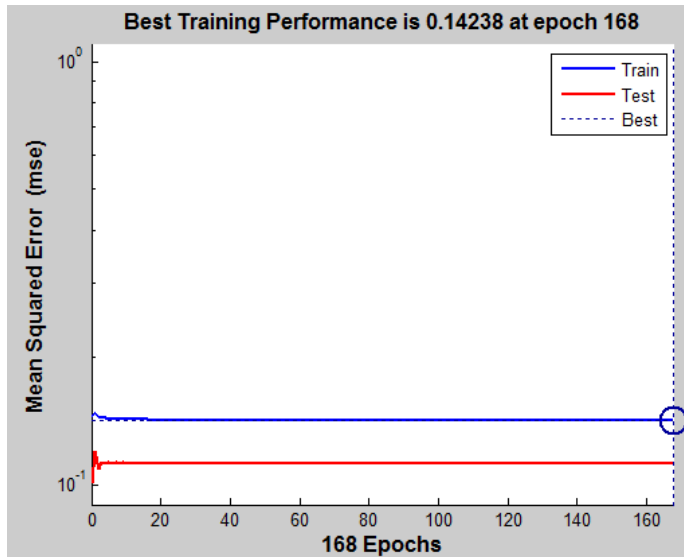
6.2.1 Implementasi 5 Variabel Input

Model paling optimal dengan 5 (lima) node input dan 1 (satu) node output didapatkan dengan struktur model yang terdiri dari 1 hidden layer yang memiliki 10 node. Fungsi aktivasi yang digunakan pada hidden layer dan output layer masing-masing adalah logsig dan purelin. Fungsi pelatihan dan pembelajaran yang digunakan adalah trainbr dan learnrgrm. Berdasarkan hasil training, model mencapai kondisi seimbang pada epoch ke 168 dengan nilai MSE sebesar 0.142. Persentase korelasi antara nilai keluaran (prediksi) dari model dengan nilai output (target) secara keseluruhan sebesar 78.39%.



Gambar 13 Struktur Model 5 Node Input dan 1 Node Output

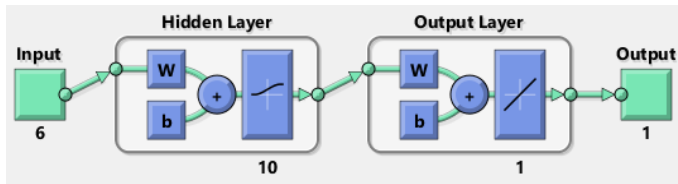
Berdasarkan grafik performa pada gambar 14, nilai error (MSE) tidak mengalami kenaikan ataupun penurunan (konstan) setelah bertambahnya jumlah epoch ketika training hingga proses training berhenti. Dari grafik tersebut dapat diperkirakan tidak akan terjadi *overfitting* pada proses training dan testing. Hal tersebut menunjukkan bahwa model yang dihasilkan akan menghasilkan performa yang baik untuk data yang dilatih saja, tetapi performanya akan berbeda ketika diuji dengan data di luar data training [17]. Berdasarkan grafik tersebut, performa data testing lebih tinggi dibandingkan data training, hal tersebut ditunjukkan dengan nilai MSE pada data training lebih tinggi daripada data testing.



Gambar 14 Grafik Performa Model 5 Node Input dan 1 Node Output

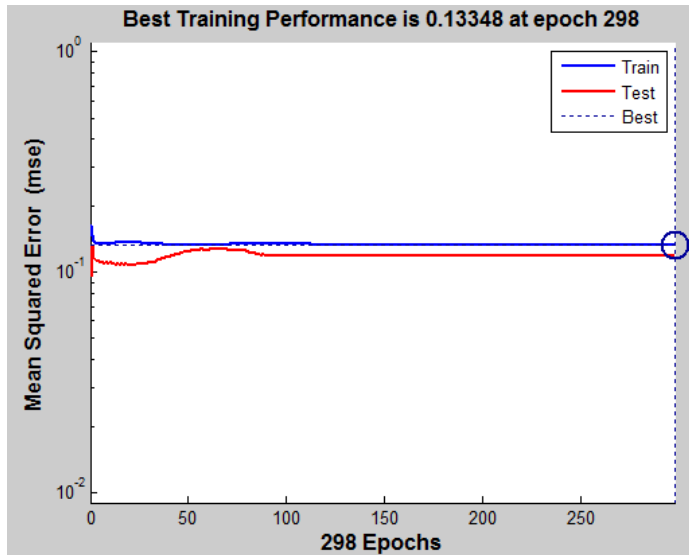
6.2.2 Implementasi 6 Variabel Input

Model paling optimal dengan 6 (enam) node input dan 1 (satu) node output didapatkan dengan struktur model yang terdiri dari 1 hidden layer yang memiliki 10 node. Fungsi aktivasi yang digunakan pada hidden layer dan output layer masing-masing adalah logsig dan purelin. Fungsi pelatihan dan pembelajaran yang digunakan adalah trainbr dan learngdm. Berdasarkan hasil training, model mencapai kondisi seimbang pada epoch ke 298 dengan nilai MSE sebesar 0.133. Persentase korelasi antara nilai keluaran (prediksi) dari model dengan nilai output (target) secara keseluruhan sebesar 79.65%.



Gambar 15 Struktur Model 6 Node Input dan 1 Node Output

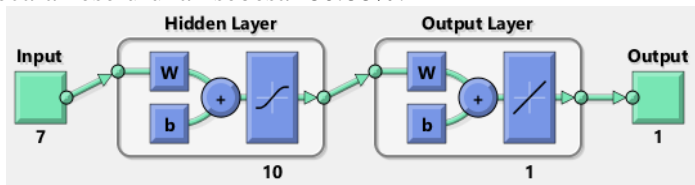
Berdasarkan grafik performa pada gambar 16, nilai error (MSE) tidak mengalami kenaikan ataupun penurunan (konstan) setelah bertambahnya jumlah epoch ketika training hingga proses training berhenti. Dari grafik tersebut dapat diperkirakan tidak akan terjadi *overfitting* pada proses training dan testing. Hal tersebut menunjukkan bahwa model yang dihasilkan akan menghasilkan performa yang baik untuk data yang dilatih saja, tetapi performanya akan berbeda ketika diuji dengan data di luar data training [17]. Berdasarkan grafik tersebut, performa data testing lebih tinggi dibandingkan data training, hal tersebut ditunjukkan dengan nilai MSE pada data training lebih rendah daripada data testing.



Gambar 16 Grafik Performa Model 6 Node Input dan 1 Node Output

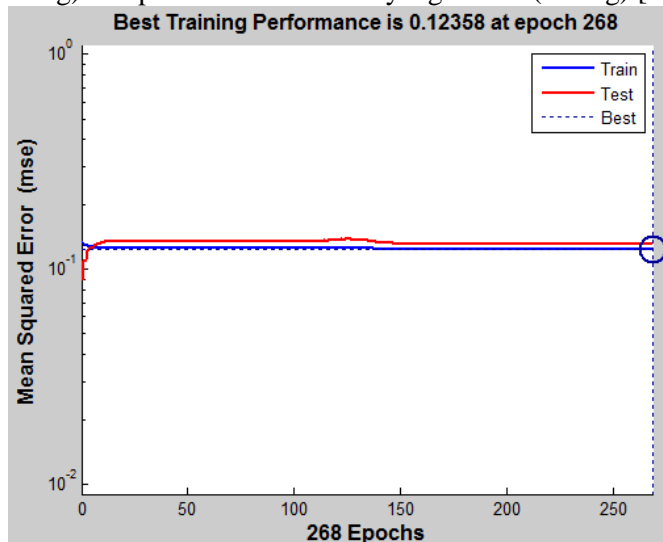
6.2.3 Implementasi 7 Variabel Input

Model paling optimal dengan 7 (tujuh) node input dan 1 (satu) node output didapatkan dengan struktur model yang terdiri dari 1 hidden layer yang memiliki 10 node. Fungsi aktivasi yang digunakan pada hidden layer dan output layer masing-masing adalah tansig dan purelin. Fungsi pelatihan dan pembelajaran yang digunakan adalah trainbr dan learnsgdm. Berdasarkan hasil training, model mencapai kondisi seimbang pada epoch ke 268 dengan nilai MSE sebesar 0.123. Persentase korelasi antara nilai keluaran (prediksi) dari model dengan nilai output (target) secara keseluruhan sebesar 80.88%.



Gambar 17 Struktur Model 7 Node Input dan 1 Node Output

Berdasarkan grafik performa pada gambar 18, nilai error (MSE) tidak mengalami kenaikan ataupun penurunan (konstan) setelah bertambahnya jumlah epoch ketika training hingga proses training berhenti. Dari grafik tersebut dapat diketahui bahwa model mengalami *overfitting* pada proses training dan testing. Hal tersebut menunjukkan bahwa model yang dihasilkan akan menghasilkan performa yang baik untuk data yang dilatih (training) maupun data di luar data yang dilatih (testing) [17].



Gambar 18 Grafik Performa Model 7 Node Input dan 1 Node Output

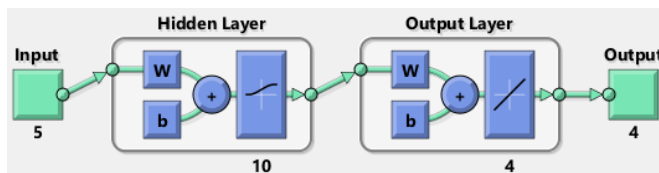
6.3 Implementasi Model 4 Node Output

Berikut ini model neural network paling optimal yang menggunakan 4 (empat) node pada output layer berdasarkan hasil training dan testing. Model tersebut dibedakan berdasarkan jumlah variabel (node) pada input layer. Pada penelitian ini dilakukan percobaan pembuatan model dengan

jumlah variabel input sebanyak 5 (lima) hingga 7 (tujuh) dari jumlah variabel input yang memiliki nilai korelasi tertinggi.

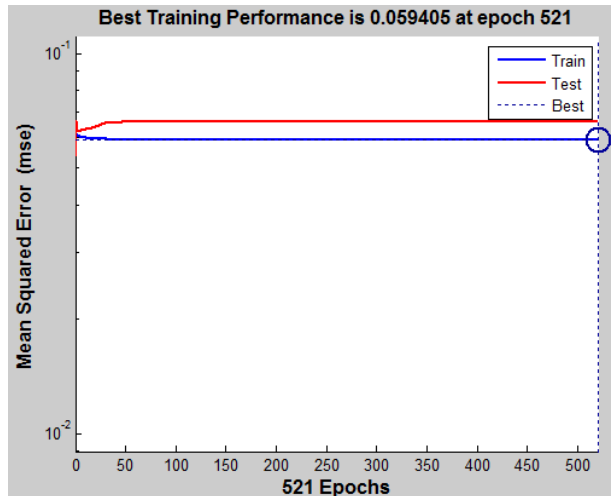
6.3.1 Implementasi 5 Variabel Input

Model paling optimal dengan 5 (lima) node input dan 4 (empat) node output didapatkan dengan struktur model yang terdiri dari 1 hidden layer yang memiliki 10 node. Fungsi aktivasi yang digunakan pada hidden layer dan output layer masing-masing adalah logsig dan purelin. Fungsi pelatihan dan pembelajaran yang digunakan adalah trainbr dan learnngdm. Berdasarkan hasil training, model mencapai kondisi seimbang pada epoch ke 521 dengan nilai MSE sebesar 0.059. Persentase korelasi antara nilai keluaran (prediksi) dari model dengan nilai output (target) secara keseluruhan sebesar 82.34%.



Gambar 19 Struktur Model 7 Node Input dan 4 Node Output

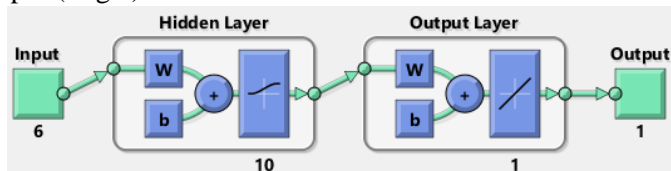
Berdasarkan grafik performa pada gambar 20, nilai error (MSE) tidak mengalami kenaikan ataupun penurunan (konstan) setelah bertambahnya jumlah epoch ketika training hingga proses training berhenti. Dari grafik tersebut dapat diperkirakan tidak akan terjadi *overfitting* pada proses training dan testing. Hal tersebut menunjukkan bahwa model yang dihasilkan akan menghasilkan performa yang baik untuk data yang dilatih saja, tetapi performanya akan berbeda ketika diuji dengan data di luar data training [17]. Berdasarkan grafik tersebut, performa data testing lebih rendah dibandingkan data training, hal tersebut ditunjukkan dengan nilai MSE pada data training lebih tinggi daripada data testing.



Gambar 20 Grafik Performa Model 7 Node Input dan 4 Node Output

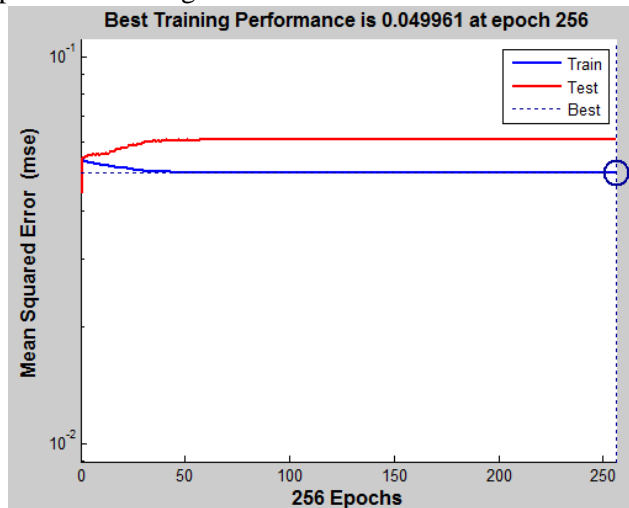
6.3.2 Implementasi 6 Variabel Input

Model paling optimal dengan 6 (enam) node input dan 4 (empat) node output didapatkan dengan struktur model yang terdiri dari 1 hidden layer yang memiliki 10 node. Fungsi aktivasi yang digunakan pada hidden layer dan output layer masing-masing adalah logsig dan purelin. Fungsi pelatihan dan pembelajaran yang digunakan adalah trainbr dan learnngdm. Berdasarkan hasil training, model mencapai kondisi seimbang pada epoch ke 256 dengan nilai MSE sebesar 0.049. Persentase korelasi antara nilai keluaran (prediksi) dari model dengan nilai output (target) secara keseluruhan sebesar 85.12%.



Gambar 21 Struktur Model 6 Node Input dan 4 Node Output

Berdasarkan grafik performa pada gambar 22, nilai error (MSE) tidak mengalami kenaikan ataupun penurunan (konstan) setelah bertambahnya jumlah epoch ketika training hingga proses training berhenti. Dari grafik tersebut dapat diperkirakan tidak akan terjadi *overfitting* pada proses training dan testing. Hal tersebut menunjukkan bahwa model yang dihasilkan akan menghasilkan performa yang baik untuk data yang dilatih saja, tetapi performanya akan berbeda ketika diuji dengan data di luar data training [17]. Berdasarkan grafik tersebut, performa data testing lebih rendah dibandingkan data training, hal tersebut ditunjukkan dengan nilai MSE pada data training lebih tinggi daripada data testing.

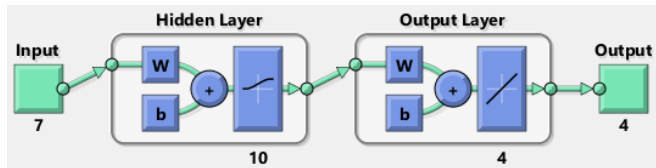


Gambar 22 Grafik Performa Model 6 Node Input dan 4 Node Output

6.3.3 Implementasi 7 Variabel Input

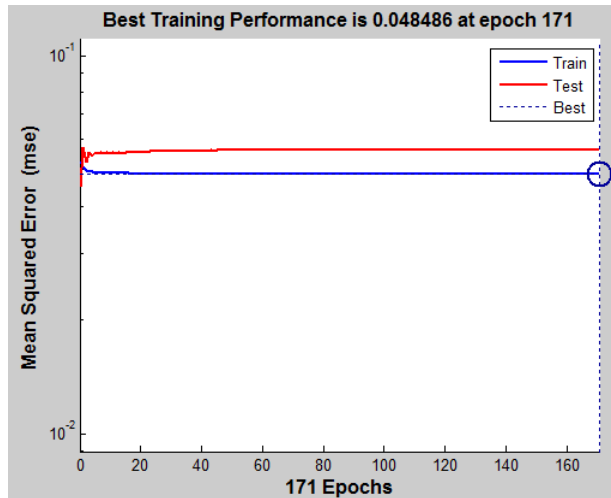
Model paling optimal dengan 7 (tujuh) node input dan 4 (empat) node output didapatkan dengan struktur model yang terdiri dari 1 hidden layer yang memiliki 10 node. Fungsi aktivasi yang digunakan pada hidden layer dan output layer masing-masing

adalah logsig dan purelin. Fungsi pelatihan dan pembelajaran yang digunakan adalah trainbr dan learngdm. Berdasarkan hasil training, model mencapai kondisi seimbang pada epoch ke 171 dengan nilai MSE sebesar 0.048. Persentase korelasi antara nilai keluaran (prediksi) dari model dengan nilai output (target) secara keseluruhan sebesar 85.75%.



Gambar 23 Struktur Model 7 Node Input dan 4 Node Output

Berdasarkan grafik performa pada gambar 24, nilai error (MSE) tidak mengalami kenaikan ataupun penurunan (konstan) setelah bertambahnya jumlah epoch ketika training hingga proses training berhenti. Dari grafik tersebut dapat diperkirakan tidak akan terjadi *overfitting* pada proses training dan testing. Hal tersebut menunjukkan bahwa model yang dihasilkan akan menghasilkan performa yang baik untuk data yang dilatih saja, tetapi performanya akan berbeda ketika diuji dengan data di luar data training [17]. Berdasarkan grafik tersebut, performa data testing lebih rendah dibandingkan data training, hal tersebut ditunjukkan dengan nilai MSE pada data training lebih tinggi daripada data testing.



Gambar 24 Grafik Performa Model 7 Node Input dan 4 Node Output

6.4 Perbandingan Recall, Precision, dan Accuracy

Tabel 50 memuat ringkasan struktur model ANN pada keenam model yang telah dilakukan uji coba.

Tabel 50 Struktur model ANN hasil percobaan

Model	A	B	C	D	E	F
Jumlah Node Input Layer	5	6	7	5	6	7
Jumlah Hidden Layer	1	1	1	1	1	1
Jumlah Node Hidden Layer	10	10	10	10	10	10
Jumlah Node Output Layer	1	1	1	4	4	4

Setelah didapatkan struktur dan parameter model ANN yang paling optimal pada setiap kombinasi jumlah node pada setiap layer, masing-masing model tersebut dilakukan uji validitas dengan data testing (data yang tidak termasuk dalam data training). Tabel 51 memuat perbandingan nilai akurasi, precision dan recall dari model yang memiliki 1 (satu) node pada output layer. Dari tabel tersebut, dapat disimpulkan bahwa

model paling optimal adalah model A, yaitu model dengan 5 (lima) variabel input (jenis kelamin, umur, berat badan, batuk, sesak). Model A memiliki nilai akurasi recall dan precision paling tinggi dibanding model B dan C.

Tabel 51 Perbandingan Recall, Precision, dan Accuracy 1 Node Output

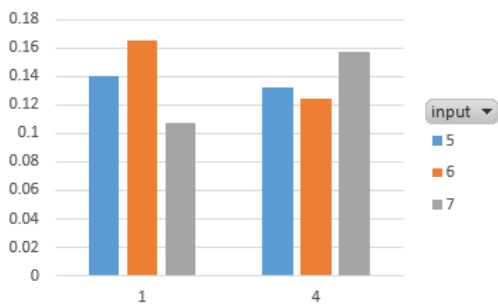
Nama Model		A	B	C
Precision (%)	PPOK	100	100	85.7
	Tuberculosis	88.5	85.56	86.67
	Bronkitis	85.18	88	95.8
	Pneumonia	#DIV/0!	0	#DIV/0!
Recall (%)	PPOK	38.89	22.22	33.33
	Tuberculosis	97.46	97.47	98.7
	Bronkitis	100	95.65	100
	Pneumonia	0	0	0
Accuracy (%)		88.43	85.12	88.43
MSE		0.140	0.165	0.107
R (%)	Training	78.28	79.08	80.84
	Test	78.6	82.9	80.83
	All	78.38	79.65	80.88

Tabel 52 memuat perbandingan nilai akurasi, precision dan recall dari model yang memiliki 4 (empat) node pada output layer. Dari tabel tersebut, dapat disimpulkan bahwa model paling optimal adalah model E, yaitu model dengan 6 variabel input (jenis kelamin, umur, berat badan, diastole, batuk, sesak). Model E memiliki nilai akurasi recall dan precision paling tinggi dibanding model D dan F.

Tabel 52 Perbandingan Recall, Precision, dan Accuracy 4 Node Output

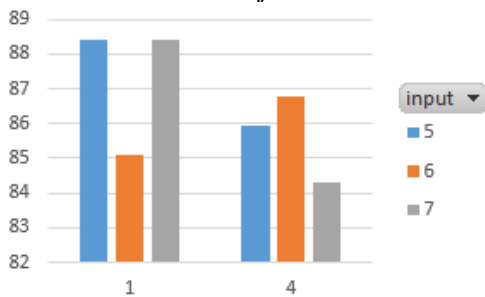
Nama Model		D	E	F
Precision (%)	PPOK	83.3	100	70
	Tuberculosis	85.39	85.56	86.7
	Bronkitis	88.46	88.46	79.3
	Pneumonia	#DIV/0!	#DIV/0!	#DIV/0!
Recall (%)	PPOK	27.78	27.78	38.89
	Tuberculosis	96.2	97.49	91.1

Nama Model		D	E	F
	Bronkitis	100	100	100
	Pneumonia	0	0	0
Accuracy (%)		85.95	86.78	84.29
MSE		0.132	0.124	0.157
R (%)	Training	82.66	85.65	86.1
	Test	80.5	82.23	84.09
	All	82.34	85.13	85.75



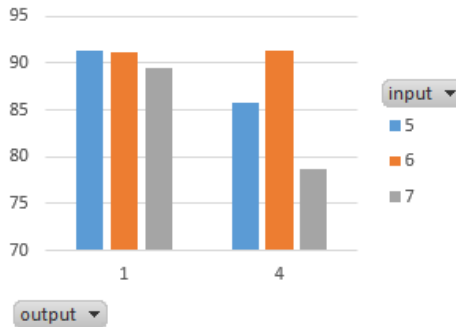
output ▼

Gambar 25 Grafik nilai MSE

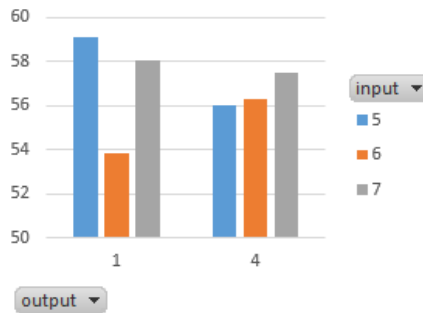


output ▼

Gambar 26 Grafik nilai akurasi



Gambar 27 Grafik nilai precision



Gambar 28 Grafik nilai recall

Dari keenam model yang telah dibuat, dapat disimpulkan bahwa tidak ada model yang dapat mengklasifikasikan jenis penyakit pneumonia. Hal tersebut ditunjukkan dari persentase nilai precision dan recall pada jenis penyakit pneumonia menghasilkan nilai 0.00%. Hal ini disebabkan oleh data training jenis penyakit pneumonia yang digunakan untuk pembuatan model sangat sedikit sehingga model ANN tidak dapat melakukan pelatihan dan pembelajaran lebih sempurna.

Selain itu, dari keenam model yang telah dibuat, dapat disimpulkan bahwa persentase nilai akurasi, precision, dan

recall tertinggi dihasilkan dari model A, yaitu model dengan jumlah node pada input layer sebanyak 5 dan jumlah node pada output layer sebanyak 1. Sedangkan nilai MSE terendah dihasilkan dari model C, yaitu model dengan jumlah node pada input layer sebanyak 7 dan jumlah node pada output layer sebanyak 1.

Pengambilan model terbaik pada penelitian ini tidak hanya berdasarkan nilai akurasi, precision, recall, dan MSE. Tetapi juga dari pertimbangan tenaga medis di RSUD Kertosono, bahwa model yang baik adalah model yang memiliki performa yang tinggi dengan jumlah variabel input yang paling banyak. Oleh sebab itu, dalam penelitian ini, model yang dipilih untuk pembuatan aplikasi berbasis desktop adalah model E, yaitu model yang mempunyai nilai precision paling tinggi dan nilai akurasi, recall, MSE yang tidak jauh dari nilai rata-rata keenam model tersebut.

6.5 Kesimpulan Hasil Percobaan

Berdasarkan hasil percobaan (training dan testing) dan uji validitas seluruh model, dapat disimpulkan bahwa pada studi kasus penelitian ini:

1. Jumlah node pada output layer tidak memberikan pengaruh yang sangat signifikan pada performa (akurasi, recall dan precision) model ANN.
2. Nilai korelasi pada variabel input mempengaruhi performa sebuah model ANN. Rata-rata model yang memiliki nilai akurasi, recall dan precision tertinggi adalah model yang menggunakan variabel input yang memiliki nilai korelasi tertinggi.
3. Jumlah data training pada masing-masing kelas (jenis penyakit) memberikan pengaruh yang sangat signifikan pada performa suatu model dalam proses klasifikasi. Semakin banyak jumlah data training pada suatu kelas,

semakin baik pula performa model tersebut dalam memprediksi kelas tersebut.

4. Jumlah hidden layer pada suatu model tidak berbanding lurus dengan performa model. Semakin banyak jumlah hidden layer pada suatu model, belum tentu menghasilkan nilai validitas yang tinggi pula.
5. Jumlah node pada hidden layer pada suatu model memberikan pengaruh yang sangat signifikan pada performa model.
6. Parameter (fungsi aktivasi, fungsi pelatihan, jumlah epoch, fungsi performa, fungsi pembelajaran) yang digunakan pada suatu model memberikan pengaruh yang sangat signifikan pada performa model.

Halaman ini sengaja dikosongkan

BAB VII

KESIMPULAN DAN SARAN

7.1 Kesimpulan

Berdasarkan penelitian dan percobaan yang telah dilakukan pada studi kasus penelitian ini, dapat disimpulkan bahwa:

Model ANN dapat digunakan untuk mendiagnosa penyakit paru dengan 6 node pada input layer (jenis kelamin, umur, berat badan, diastole, batuk, sesak), 10 node pada 1 hidden layer, dan 4 node pada output layer (ppok, tuberculosis, bronkitis, pneumonia) dengan nilai akurasi sebesar 86.78%.

7.2 Saran

Berdasarkan penelitian dan percobaan yang telah dilakukan pada studi kasus penelitian ini, terdapat beberapa saran dari peneliti terkait pembuatan model klasifikasi dengan metode ANN. Saran ini bertujuan untuk meningkatkan performa dan validitas model yang akan dikembangkan, antara lain:

1. Data sample yang digunakan tidak hanya bersumber dari RSUD Kertosono, tetapi juga ditambahkan dengan data dari instansi kesehatan (puskesmas atau rumah sakit) lainnya.
2. Data yang digunakan tidak hanya dari data pasien rawat jalan, tetapi juga dari data pasien rawat inap.
3. Menambah data dari periode sebelum Januari 2014 dan setelah Desember 2016.
4. Menambah jenis penyakit paru yang akan diklasifikasikan.

Halaman ini sengaja dikosongkan

DAFTAR PUSTAKA

- [1] WHO. [Online]. Available: <http://www.who.int/features/factfiles/>. [Accessed: 01-Feb-2017].
- [2] O. Er, N. Yumusak, and F. Temurtas, "Chest diseases diagnosis using artificial neural networks," *Expert Systems with Applications*, vol. 37, no. 12, pp. 7648–7655, Dec. 2010.
- [3] *Sistem Informasi Manajemen 2 (ed.10)*. Penerbit Salemba.
- [4] Laurene Fausett, 1993, *Fundamental of Neural Network: Architecture, Algorithm and Applications*.
- [5] O. Andareto and P. I. Semesta, *Penyakit Menular di Sekitar Anda: mengetahui macam-macam penyakit yang dapat menular serta cara pencegahannya*. Lembar Langit Indonesia, 2015.
- [6] *Farmakologi*. EGC.
- [7] *IPA Terpadu SMP/MTs Kls IX A*. Grasindo.
- [8] Y. Wu *et al.*, "An optimal tumor marker group-coupled artificial neural network for diagnosis of lung cancer," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11329–11334, Sep. 2011.
- [9] J. B. de O. Souza Filho *et al.*, "A screening system for smear-negative pulmonary tuberculosis using artificial neural networks," *International Journal of Infectious Diseases*, vol. 49, pp. 33–39, Aug. 2016.
- [10] Martin T. Hagan, dkk. *Neural Network Design, 2nd Edition*.
- [11] Flach, Peter. 2012. *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*. Cambridge University Press.
- [12] Kusrini, E. taufiq luthfi, and S. Amikom, *algoritma data mining*. Penerbit Andi.
- [13] F. Amato, A. López, E. M. Peña-Méndez, P. Vañhara, A. Hampl, and J. Havel, "Artificial neural networks in medical diagnosis," *Journal of Applied Biomedicine*, vol. 11, no. 2, pp. 47–58, 2013.

- [14] R. K. Bansal, A. Goel, and M. K. Sharma, *MATLAB and Its Applications in Engineering*. Pearson Education India, 2009.
- [15] *Bs Database Dgn Delphi 7*. Elex Media Komputindo
- [16] "MathWorks - Makers of MATLAB and Simulink." [Online]. Available: <https://www.mathworks.com/>
- [17] H. Gumilang, *Sistem Trading Kuantitatif: Membangun dan Menguji Sistem Trading Saham yang Tangguh, Efektif, dan bisa Diandalkan dalam menghasilkan CUAN di Bursa SAham*. HGU Publishing, 2013.
- [18] *Teknik2 Statistik dalam Bisnis dan Ekonomi 2 (ed.13)*. Penerbit Salemba.
- [19] *Riset Keuangan Pengujian-pengujian empiris*. Gramedia Pustaka Utama.
- [20] R. K. Bansal, A. Goel, and M. K. Sharma, *MATLAB and Its Applications in Engineering*. Pearson Education India, 2009.
- [21] A. S. Hamdi and E. Bahrudin, *Metode Penelitian Kuantitatif: Aplikasi dalam Pendidikan*. Deepublish, 2014.
- [22] *Statistik untuk Kedokteran dan Kesehatan*. Penerbit Salemba.
- [23] Santoso, *Statistika Hospitalitas*. Deepublish, 2016.
- [24] *Panduan Lab Statistik Inferensial*. Grasindo.
- [25] J. F. Steffensen, *Interpolation: Second Edition*. Courier Corporation, 2012.

BIODATA PENULIS



Penulis lahir di Nganjuk, 21 November 1995, dengan nama lengkap Nurita Damayanti. Penulis merupakan anak pertama dari tiga bersaudara.

Riwayat pendidikan penulis yaitu TK Yos Sudarso, SD Negeri Kutorejo 2 Kertosono, SMP Negeri 2 Kertosono, SMA Negeri 1 Kertosono, dan akhirnya menjadi salah satu mahasiswa Sistem Informasi angkatan 2013 melalui jalur SNMPTN Undangan dengan NRP 5213-100-083.

Selama kuliah, penulis pernah aktif di BEM FTIf sebagai staf Student Resource Development pada tahun 2015-2016, di Lembaga Dakwah Jurusan - Kajian Islam Sistem Informasi (KISI) sebagai Ketua Keputrian pada tahun 2015-2016, serta di Jama'ah Masjid Manarul Ilmi (JMMI) – TPPI ITS sebagai Wakil Direktur Badan Pelayanan Ummat (BPU) pada tahun 2016-2017. Penulis mengambil bidang minat Rekayasa Data dan Intelegensi Bisnis. Penulis dapat dihubungi melalui email nuritad3@gmail.com.

LAMPIRAN A
DATA MENTAH (7 VARIABEL AWAL)

Tabel 53 Data mentah penyakit ppok (7 variabel)

No.	Jenis Kelamin	Umur	Berat Badan	Sistole	Diastole	Batuk	Sesak
1	0	60	51	105	67	1	1
2	1	76	27	120	80	1	1
3	0	79	27.29	120	70	1	1
4	0	77	27.58	174	81	1	1
5	0	75	27.87	125	70	1	1
6	0	63	28.16	140	80	1	1
7	0	74	28.44	114	71	1	0
8	0	74	28.73	130	70	1	1
9	0	68	29.02	171	86	1	0
10	0	80	29.31	129	71	0	0
11	0	84	29.6	141	85	1	1
12	0	55	29.89	122	70	1	1
13	1	80	30.18	136	72	0	0
14	0	79	30.47	140	81	0	0
15	0	57	30.76	117	71	1	1
16	0	51	31.04	133	84	1	1
17	1	68	31.33	170	90	1	1
18	0	66	31.62	190	157	0	1
19	0	73	31.91	142	83	0	1
20	1	55	32.2	160	87	1	1
21	0	74	32.49	105	70	1	1
22	1	60	32.78	140	80	1	1
23	1	68	33.07	128	58	1	1

24	0	70	33.36	155	84	1	1
25	1	71	33.64	148	79	1	1
26	1	74	33.93	126	70	1	1
27	1	65	34.22	140	70	1	1
28	0	63	34.51	130	80	1	1
29	0	76	34.8	121	187	1	1
30	0	73	35.09	148	94	1	1
31	0	56	35.38	143	79	1	1
32	0	56	35.67	100	70	1	1
33	1	42	35.96	94	63	1	1
34	0	54	36.24	116	80	0	0
35	1	52	36.53	120	70	1	1
36	0	68	36.82	110	70	1	1
37	0	50	37.11	151	95	1	1
38	0	55	37.4	120	70	2	1
39	0	74	37.69	117	64	1	1
40	0	69	37.98	110	70	1	1
41	0	64	38.27	113	73	1	1
42	1	69	38.56	110	58	1	0
43	1	61	38.84	150	90	1	1
44	0	60	39.13	130	70	1	1
45	1	56	39.42	120	90	0	1
46	0	64	39.71	130	80	1	1
47	0	68	40	124	73	1	1
48	0	43	50	102	63	1	0
49	0	62	47	162	97	0	1
50	1	93	33	110	50	1	1
51	0	83	32	100	70	1	0

52	0	73	31	110	70	0	1
53	0	70	30	100	66	1	1
54	0	69	29	120	70	0	1
55	0	61	28	107	67	0	0
56	0	80	27	106	44	1	1
57	1	21	26	90	60	0	0
58	1	13	25	92	60	1	1
59	0	60	26.83	131	73	1	1
60	1	65	28.67	130	80	1	1
61	0	69	30.5	120	60	1	1
62	1	35	32.33	120	70	0	1
63	1	30	34.17	120	70	0	0
64	0	56	36	115	80	1	1
65	1	74	37	120	70	1	1
66	0	51	40	120	80	1	1
67	1	15	43	107	69	0	0
68	1	50	42.6	168	98	1	1
69	1	79	42.2	158	74	1	1
70	0	69	41.8	90	70	1	
71	0	50	41.4	140	80	1	0
72	1	17	41	110	70	1	0

Tabel 54 Data mentah penyakit tuberculosis (7 variabel)

No.	Jenis Kelamin	Umur	Berat Badan	Sistole	Diastole	Batuk	Sesak
1	0	73	40.5	159	101	2	0
2	0	73	40	100	60	1	1
3	0	60	39.5	153	105	0	1
4	0	72	39	90	60	2	1
5	1	43	32	136	104	2	1

6	0	39	40	104	74	0	1
7	0	33	60	106	69	2	0
8	1	81	33	135	79	0	0
9	1	50	36	120	53	1	0
10	1	48	43	121	79	2	0
11	1	49	52	111	68	1	0
12	0	61	31	130	72	0	0
13	0	69	29	96	63	0	1
14	1	61	38	112	64	1	0
15	1	35	49	93	61	2	0
16	0	60	49	116	70	1	0
17	1	50	49	113	62	1	0
18	1	56	46	124	68	0	0
19	1	63	43	171	86	1	0
20	1	74	39	101	62	0	0
21	1	36	50	121	75	0	0
22	0	68	39	97	66	1	1
23	0	64	39	115	83	2	1
24	1	26	40	120	70	1	0
25	1	58	58	130	80	1	0
26	1	50	59	166	130	1	0
27	1	45	55	139	82	1	0
28	0	65	41.5	176	99	1	0
29	1	42	29	138	87	2	0
30	0	62	30	130	80	1	0
31	1	50	31	140	90	1	0
32	0	53	32	127	75.5	1	1
33	0	83	33	114	61	1	0

34	0	52	51	116	69	2	0
35	0	70	46.5	160	111	1	0
36	0	54	42	104	67	2	1
37	0	59	41	128	81	1	1
38	0	75	40	130	78	1	1
39	0	14	39	96	63	2	0
40	1	60	26	129	68	1	0
41	1	42	53	149	91	1	1
42	0	21	49	120	78	1	1
43	0	53	45	124	82	1	1
44	0	61	48	120	70	2	0
45	0	27	47.5	117	64	2	0
46	1	17	51	87	57	1	0
47	0	47	38	96	64	1	1
48	1	61	25	105	71	1	0
49	0	78	31.5	135	85	1	0
50	0	25	38	96	62	1	1
51	0	43	56	90	57	2	0
52	1	18	48	110	54	1	0
53	0	78	31	82	57	1	1
54	0	33	45.5	87	59	0	1
55	1	81	33	131	74	1	1
56	1	55	40	125	74	2	0
57	0	80	29	100	60	1	0
58	0	59	60	137	83	1	1
59	1	45	35	140	80	2	0
60	0	48	43	101	66	1	1
61	0	38	35	96	67	1	1

62	1	70	27	120	70	1	1
63	1	65	28	119	71	1	0
64	1	65	34.75	96	55	1	0
65	1	55	41.5	142	110	0	0
66	1	68	48.25	111	74	1	1
67	1	50	55	142	81	1	1
68	0	33	67	107	71	1	1
69	1	49	54.5	90	60	1	1
70	0	76	42	128	69	2	0
71	1	20	41	110	69	1	0
72	0	36	45	129	83	1	0
73	1	64	51	124	62	1	0
74	0	40	48	125	91	1	0
75	1	28	33	93	59	1	1
76	0	30	61	106	69	1	0
77	1	39	40	115	79	2	0
78	1	38	39	90	60	1	0
79	0	68	45	108	72	1	0
80	1	67	26	109	71	2	0
81	0	6	41	120	70	1	0
82	1	37	39	113	60	1	0
83	1	34	27.5	110	87	1	1
84	1	60	27	110	70	1	0
85	1	50	46.5	110	70	1	0
86	1	45	55	122	83	0	0
87	1	74	37	124	105	0	0
88	1	80	26	151	89	0	0
89	0	49	55	114	70	1	0

90	1	64	42	104	57	2	0
91	0	59	51	120	68	2	0
92	1	31	43	96	62	1	1
93	1	57	55	118	74	2	0
94	0	79	32	124	78	2	0
95	1	40	35	106	76	1	0
96	0	70	38	99	63	2	0
97	1	72	51	153	83	2	1
98	0	54	43	113	70	2	0
99	1	65	35	114	70	0	0
100	0	72	42.25	83	48	1	1
101	0	80	49.5	105	62	1	0
102	0	75	32	127	76	1	0
103	1	40	40	115	80	1	0
104	1	85	43	131	65	2	0
105	0	41	46	77	54	1	1
106	0	42	56	128	69	1	0
107	0	54	55	126	69	1	1
108	0	31	52	110	70	1	0
109	0	29	50	137	89	1	0
110	0	42	44	92	57	2	1
111	1	50	46	143	78	1	0
112	0	65	46	133	79	1	0
113	0	35	57.33	98	64	2	0
114	1	76	68.67	108	76	1	0
115	0	51	80	124	83	1	1
116	0	21	62	114	69	1	0
117	1	52	44	93	64	1	0

118	0	39	57	113	74	1	0
119	0	54	58	134	85	0	0
120	0	76	46	132	78	2	0
121	0	51	50	111	66	2	1
122	1	50	49.33	112	67	1	0
123	0	60	48.67	120	70	1	0
124	0	71	48	134	72	1	1
125	0	57	62	123	70	0	0
126	1	49	35	93	56	1	0
127	1	37	36	153	96	1	0
128	0	55	42	99	66	1	1
129	1	44	34	141	82	2	1
130	0	37	54	93	58	1	1
131	0	35	43	123	82	0	0
132	1	23	47	127	90	1	1
133	1	34	41	95	63	1	1
134	1	59	40	122	68	2	0
135	0	65	39.75	90	70	2	1
136	1	40	39.5	170	100	1	0
137	0	56.5	39.25	120	70	0	0
138	0	73	39	110	60	1	0
139	0	71	36.25	95	67	2	1
140	0	24	33.5	120	75	1	0
141	1	55	30.75	100	70	0	0
142	1	49	28	140	80	2	1
143	0	28	44	110	70	1	1
144	1	58	48.5	116.5	80.5	0	0
145	0	45	53	123	91	1	1

146	0	39	45	110	70	1	1
147	0	32	43.5	100	70	1	0
148	1	55	42	110	70	1	0
149	0	19	65	120	70	0	0
150	1	43	55	155	100	1	1
151	1	70	42	150	80	1	0
152	0	42	38	85	60	1	1
153	0	49	39	115	77	1	0
154	0	20	43	120	80	1	0
155	0	75	43.67	90	60	1	1
156	0	40	44.33	120	70	30	1
157	0	29	45	100	60	1	0
158	1	27	44.67	100	70	2	1
159	0	32	44.33	80	60	1	1
160	0	60	44	130	80	1	1
161	1	55	47.33	130	70	2	0
162	0	80	50.67	170	110	0	0
163	1	25	54	130	90	1	1
164	0	43	44.5	90	60	1	1
165	0	58	35	145	96	1	0
166	1	40	35	120	70	1	1
167	0	43	46	110	70	1	1
168	1	57	65	115	67	2	0
169	1	69	58.5	165	81	2	0
170	0	47	52	100	70	1	1
171	1	76	49.33	140	90	1	1
172	0	80	46.67	175	90	1	1
173	1	50	44	120	80	1	1

174	1	36	44	100	80	2	0
175	0	51	44	140	80	1	1
176	1	67	44	117	73	1	0
177	1	57	48.67	128.5	76.5	1	1
178	1	42	53.33	140	80	1	0
179	0	67	58	130	80	1	0
180	1	55	56.83	180	100	2	0
181	0	33	55.67	126	70	1	0
182	1	41.5	54.5	130	80	0	0
183	1	50	53.33	160	100	1	1
184	1	60	52.17	120	70	1	1
185	0	25	51	115	74	1	0
186	1	57	48	102	63	1	0
187	1	64	48	140	80	0	1
188	0	66	48	127	82	0	0
189	0	41	48	120	70	1	0
190	0	46	51	130	80	0	0
191	0	54	54	100	70	1	1
192	1	70	51	120	70	1	1
193	1	69	48	134	63	0	0
194	1	22	45	124	78	2	1
195	1	31	45	124	93	1	1
196	0	36	45	130	90	2	0
197	1	22	32	87	58	1	0
198	1	55	37	130	80	2	0
199	0	57	35.6	160	90	1	1
200	0	55	34.2	105	60	1	1
201	0	55	32.8	118	73	1	0

202	0	60	31.4	78	50	1	1
203	0	70	30	108	59	0	0
204	1	55	31	107	57	1	1
205	0	30	32	95	65	0	0
206	1	49	33	124	79	2	0
207	1	35	34.13	155	100	0	0
208	0	73	35.25	120	80	1	1
209	0	73	36.38	130	70	0	1
210	0	71	37.5	120	70	1	1
211	0	82	38.63	180	110	1	1
212	0	42	39.75	104	70	0	0
213	0	13	40.88	123	75	0	0
214	1	66	42	111	56	1	0
215	1	35	38.5	101	66	1	0
216	1	21	35	99	64	1	0
217	0	57	37.5	93	65	1	1
218	0	65	40	140	86	1	0
219	0	51	40.4	85	47	1	1
220	1	65	40.8	129	98	1	0
221	1	58	41.2	156	97	2	0
222	0	52	41.6	124	93	1	0
223	0	33	42	99	64	1	0
224	0	62	43.5	120	74	1	1
225	0	14	45	115	82	1	0
226	0	65	46.5	125	85	1	0
227	0	35	48	114	73	1	1
228	0	71	44	128	81	1	1
229	1	85	30	140	80	1	1

230	1	42	57	147	82	0	0
231	1	68	55.75	121	49	1	1
232	1	38	54.5	95	66	1	0
233	1	76	53.25	140	82	1	1
234	1	53	52	106	68	1	1
235	0	60	50.75	130	70	1	0
236	0	40	49.5	100	70	2	0
237	1	30	48.25	106	61	1	0
238	1	34	47	130	60	1	1
239	0	46	45.33	140	80	2	0
240	1	70	43.67	135	68	2	0
241	0	60	42	90	60	1	1
242	0	65	45	117	60	1	0
243	1	65	57	109	90	1	1
244	1	44	69	128	85	1	1
245	1	37	66.75	100	60	2	0
246	1	57	64.5	116	76	2	1
247	1	27	62.25	97	70	1	0
248	0	53	60	114	82.7	0	1
249	1	33	37	131	95.3	1	0
250	1	38	39.5	148	108	0	0
251	0	43	42	151	78	2	0
252	0	60	44.5	160	90	0	0
253	0	66	47	110	54	2	0
254	0	51	49.5	80	60	1	1
255	0	70	52	140	75	2	1
256	1	45	51.87	120	80	1	1
257	0	43	51.73	127	80	2	1

258	0	51	51.6	139	92	1	0
259	1	41	51.47	150	45	1	1
260	0	30	51.33	147	93	2	1
261	1	31	51.2	100	60	2	0
262	0	26	51.07	135	70	1	0
263	1	31	50.93	109	77	1	0
264	1	36	50.8	137	88	0	0
265	1	41	50.67	120	70	1	1
266	1	59	50.53	123	73	1	0
267	0	51	50.4	123	83	1	1
268	0	52	50.27	81	66	1	1
269	0	70	50.13	170	90	1	1
270	0	26	50	110	70	1	1
271	1	58	46.67	110	73	1	1
272	0	33	43.33	100	70	1	1
273	0	30	40	94	62	1	1
274	0	33	38	121	79	1	0
275	0	19	44	85	59	0	0
276	0	66	54	120	70	1	1
277	0	40	52	93	65	1	0
278	0	35	62	118	62	1	1
279	0	18	65	114	65	0	0
280	0	32	32	121	81	1	0
281	0	39	47	100	70	1	1
282	0	38	30	110	60	1	1
283	1	33	35	120	75	1	1
284	0	28	50	120	70	1	0
285	1	55	49	120	70	1	1

286	0	37	48	100	70	1	1
287	0	48	57	120	80	0	0
288	1	53	57.5	90	70	0	0
289	1	42	58	115	80	1	0
290	1	37	58.67	112.5	75	1	0
291	0	40	59.33	110	70	1	1
292	0	52	60	100	60	1	1
293	0	60	41	150	80	2	0
294	0	16	42.04	100	70	2	1
295	1	75	43.08	160	90	1	1
296	1	47	44.12	168	78	1	1
297	0	50	45.15	120	70	1	1
298	1	59	46.19	131	61	2	0
299	1	45	47.23	170	90	1	0
300	0	51	48.27	107	73	1	1
301	0	44	49.31	110	70	1	0
302	1	48	50.35	92	47	1	0
303	0	50	51.38	110	75	2	1
304	1	46	52.42	125	50	2	0
305	1	56	53.46	130	80	1	1
306	1	77	54.5	150	130	1	1
307	0	54	55.54	140	90	2	0
308	1	45	56.58	130	70	1	0
309	0	55	57.62	138	89	1	0
310	0	50	58.65	170	110	2	0
311	1	40	59.69	125	72	2	0
312	1	48	60.73	130	80	1	0
313	1	19	61.77	112	77	2	1

314	0	62	62.81	152	88	2	0
315	0	65	63.85	121	77	1	1
316	0	67	64.88	154	94	0	0

Tabel 55 Data mentah penyakit bronkitis (7 variabel)

No.	Jenis Kelamin	Umur	Berat Badan	Sistole	Diastole	Batuk	Sesak
1	0	69	65.92	154	133	1	1
2	0	69	66.96	120	70	1	1
3	1	48	68	141	77	1	1
4	1	58	52	160	92	1	0
5	1	31	48.5	121	81	2	0
6	0	20	45	130	70	1	0
7	1	50	32	130	80	1	0
8	1	84	25	130	80	1	1
9	0	68	29	110	70	1	1
10	1	40	33	90	60	2	1
11	1	43	58	250	130	2	0
12	0	36	49	140	80	2	1
13	0	40	39	102	69	1	1
14	1	78	29	130	80	1	0
15	1	59	55	145	80	1	0
16	0	59	44	111	65	1	0
17	0	39	33	110	80	1	1
18	1	48	37.75	100	50	2	0
19	0	58	42.5	143	89	1	1
20	1	65	47.25	140	90	1	0
21	0	72	52	131	69	1	1
22	0	34	44.67	84	53	1	0
23	1	52	37.33	135	80	2	0

24	1	34	30	175	110	0	1
25	0	18	60	100	70	1	0
26	1	72	42	130	70	1	0
27	1	54	54	98	68	1	0
28	1	56	36	120	51	0	0
29	1	34	47	150	90	1	1
30	0	73	39	130	70	1	0
31	1	36	49	110	70	1	1
32	1	48	68	146	81	1	0
33	0	72	42	160	96	1	1
34	1	67	58	134	64	0	0
35	1	64	42	141.5	72	0	0
36	1	56	36	149	80	1	0
37	1	59	44	166	94	1	0
38	0	37	48	130	85	0	0
39	0	18	60	110	68	1	0
40	0	71	44	128	74	0	0
41	1	77	50	76	46	0	0
42	1	21	35	103	70	0	0
43	1	80	49.5	99	69	1	1
44	1	68	39	113	55	1	0
45	0	52	60	104	67	0	0
46	1	32	32	95	63	1	0
47	1	51	50	128	76	1	0
48	1	24	43	104	72	0	0
49	1	50	32	136	82	0	1
50	1	47	52	145	88	1	1
51	1	67	58	123	87	0	0

52	0	53	60	112	83	1	0
53	0	41	47	100	70	1	1
54	1	62	47	150	80	1	1
55	0	55	37	130	70	1	1
56	1	38	30	110	70	0	0
57	0	51	50	144	88	1	1
58	0	79	29	145	70	1	1
59	0	74	39	110	70	1	0
60	1	29	45	99	56	1	0
61	0	74	39	120	65	1	1
62	1	34	30	145	96	1	0
63	0	24	43	94	63	1	0
64	1	76	46	109	74	1	0
65	0	38	30	130	70	1	1
66	1	64	42	131	71	1	0
67	0	13	25	120	70	1	1
68	1	61	47	119	80	1	1
69	1	65	45	117	70	1	0
70	1	39	45	120	76	1	1
71	0	51	50	127	71.5	1	0
72	0	63	43	134	67	1	1
73	1	56	36	104	65	1	1
74	0	40	33	120	80	0	0
75	1	36	49	131	77	1	0
76	1	52	60	109	66	1	1
77	1	78	29	150	80	1	1
78	1	31	52	155	80	0	0
79	0	71	44	160	80	1	0

80	0	74	39	150	80	1	0
81	0	49	39	124	66	0	1
82	1	58	52	110	73	1	1
83	1	33	35	140	80	1	0
84	0	71	44	102	73	0	0
85	1	51	50	108	63	0	0
86	1	61	47	110.7	62.3	1	1
87	0	52	60	113.3	61.7	1	0
88	1	71	44	116	61	1	0
89	0	60	41	114	79	1	1
90	0	21	35	115	75	1	1
91	1	56	36	128	90	1	0
92	0	62	47	120	70	1	0

Tabel 56 Data mentah penyakit pneumonia (7 variabel)

No.	Jenis Kelamin	Umur	Berat Badan	Sistole	Diastole	Batuk	Sesak
1	1	63		107	61	1	1
2	0	60		110	70	0	0
3	1	67		131	77	1	0

LAMPIRAN B

TABEL KONTINGENSI

Tabel 57 Tabel Kontingensi 5 Node Input 1 Node Output

	Prediksi PPOK	Prediksi Tb	Prediksi Bronkitis	Prediksi Pneumonia	Recall
Aktual PPOK	6	11	1	0	0.33
Aktual Tb	0	77	2	0	0.97
Aktual Bronkitis	0	0	23	0	1
Aktual Pneumonia	0	0	1	0	0
Precision	1	0.875	0.85	#DIV/0!	
	Accuracy		0.876		

Tabel 58 Tabel Kontingensi 6 Node Input 1 Node Output

	Prediksi PPOK	Prediksi Tb	Prediksi Bronkitis	Prediksi Pneumonia	Recall
Aktual PPOK	4	13	1		0.22
Aktual Tb	0	77	2	0	0.975
Aktual Bronkitis	0	0	22	1	0.956
Aktual Pneumonia	0	0	0	0	0
Precision	1	0.856	0.88	0	
	Accuracy		0.851		

Tabel 59 Tabel Kontingensi 7 Node Input 1 Node Output

	Prediksi PPOK	Prediksi Tb	Prediksi Bronkitis	Prediksi Pneumonia	Recall
Aktual PPOK	6	12	0	0	0.33
Aktual Tb	1	78	0	0	0.987
Aktual Bronkitis	0	0	23	0	1
Aktual Pneumonia	0	0	1	0	0
Precision	0.857	0.867	0.958	#DIV/0!	

Accuracy	0.8843
----------	--------

Tabel 60 Tabel Kontingensi 5 Node Input 4 Node Output

	Prediksi PPOK	Prediksi Tb	Prediksi Bronkitis	Prediksi Pneumonia	Recall
Aktual PPOK	5	13	0	0	0.278
Aktual Tb	1	76	2	0	0.962
Aktual Bronkitis	0	0	23	0	1
Aktual Pneumonia	0	0	1	0	0
Precision	0.833	0.854	0.885	#DIV/0!	
Accuracy		0.859			

Tabel 61 Tabel Kontingensi 6 Node Input 4 Node Output

	Prediksi PPOK	Prediksi Tb	Prediksi Bronkitis	Prediksi Pneumonia	Recall
Aktual PPOK	5	13	0	0	0.278
Aktual Tb	0	77	2	0	0.975
Aktual Bronkitis	0	0	23	0	1
Aktual Pneumonia	0	0	1	0	0
Precision	1	0.856	0.885	#DIV/0!	
Accuracy		0.867			

Tabel 62 Tabel Kontingensi 7 Node Input 4 Node Output

	Prediksi PPOK	Prediksi Tb	Prediksi Bronkitis	Prediksi Pneumonia	Recall
Aktual PPOK	7	11	0	0	0.389
Aktual Tb	3	72	5	0	0.911
Aktual Bronkitis	0	0	23	0	1
Aktual Pneumonia	0	0	1	0	0
Precision	0.7	0.868	0.793	#DIV/0!	
Accuracy		0.843			

LAMPIRAN C

DOKUMENTASI APLIKASI

Lung Patient Confirmation

Patient's diagnosis

Jenis Kelamin	Laki-laki	<div>train</div> <div>Diagnose</div> <div>tuberculosis</div>	ppok	7.94966 %
Umur	50		tuberculosis	77.9521 %
Berat Badan	45		bronkitis	14.2765 %
Diastole	110		pneumonia	0.177443 %
Batuk	Tidak Batuk			
Sesak	Tidak			

Gambar 29 Tampilan Aplikasi di Matlab